

# Life in the Atacama: Science autonomy for improving data quality

Trey Smith,<sup>1</sup> David R. Thompson,<sup>1</sup> David S. Wettergreen,<sup>1</sup> Nathalie A. Cabrol,<sup>2</sup> Kimberley A. Warren-Rhodes,<sup>2</sup> and Shmuel J. Weinstein<sup>3</sup>

Received 19 September 2006; revised 1 April 2007; accepted 6 August 2007; published 7 December 2007.

[1] “Science autonomy” refers to exploration robotics technologies involving onboard science analysis of collected data. These techniques enable a rover to make adaptive decisions about which measurements to collect and transmit. Science autonomy can compensate for limited communications bandwidth by ensuring that planetary scientists receive those images and spectra that best meet mission goals. Here, we present the results of autonomous science experiments performed in the Atacama Desert of Chile during the Life in the Atacama (LITA) rover field campaign. We aim to provide an overview of autonomous science principles and examine their integration into the LITA operations strategy. We present experiments in four specific autonomous science domains: (1) autonomously responding to evidence of life with more detailed measurements; (2) rock detection for site profiling and selective data return; (3) tactical replanning to efficiently map the distribution of life; (4) detecting novel images and geologic unit boundaries in image sequences. In each of these domains we demonstrate improvements in the quality of returned data through autonomous analysis of imagery.

**Citation:** Smith, T., D. R. Thompson, D. S. Wettergreen, N. A. Cabrol, K. A. Warren-Rhodes, and S. J. Weinstein (2007), Life in the Atacama: Science autonomy for improving data quality, *J. Geophys. Res.*, 112, G04S03, doi:10.1029/2006JG000315.

## 1. Introduction

[2] Communication with planetary rovers is characterized by low bandwidth and long latency. Rovers’ increasing speed and mobility will intensify this communication bottleneck. Future generations of rovers will travel long distances per command cycle; D. Wettergreen et al. (unpublished manuscript, 2006) regularly demonstrated traverses over five kilometers. These “over-the-horizon” traverses permit broad surveys of regions containing multiple geologic and environmental units. However, scientists planning segments of the traverse that are over the horizon must rely on orbital imagery alone to schedule data collection actions. Moreover, bandwidth constraints prevent the return of detailed measurements covering the entire traverse. Planned bandwidth upgrades to the Deep Space Network will not keep pace with improvements in rover mobility and navigation; this risks situations where most of the terrain visited by the rover is never seen by scientists. Thus the current strategy of sending an explicit command sequence for each measurement scales poorly to long-range rovers.

[3] “Science autonomy” refers to exploration robotics technologies involving onboard science analysis of collect-

ed data [Gilmore et al., 2000; Gulick et al., 2001; Castaño et al., 2003; Smith et al., 2005]. This onboard analysis addresses the communication bottleneck in several ways. First, it permits the rover to make immediate, adaptive data collection decisions in response to its own observations. In this way it can direct finite time and energy resources toward observing the most significant features. Rovers employing adaptive data collection can obviate the need for return trips in later command cycles. Return trips are particularly costly because of accumulated localization error; returning to a single specific target can easily require multiple command cycles.

[4] In addition to adaptive sampling, spacecraft equipped with onboard data understanding can help prioritize data for return. Scientists can schedule dense measurements—far more data than bandwidth can accommodate—and command the science-aware spacecraft to identify the most relevant data to downlink.

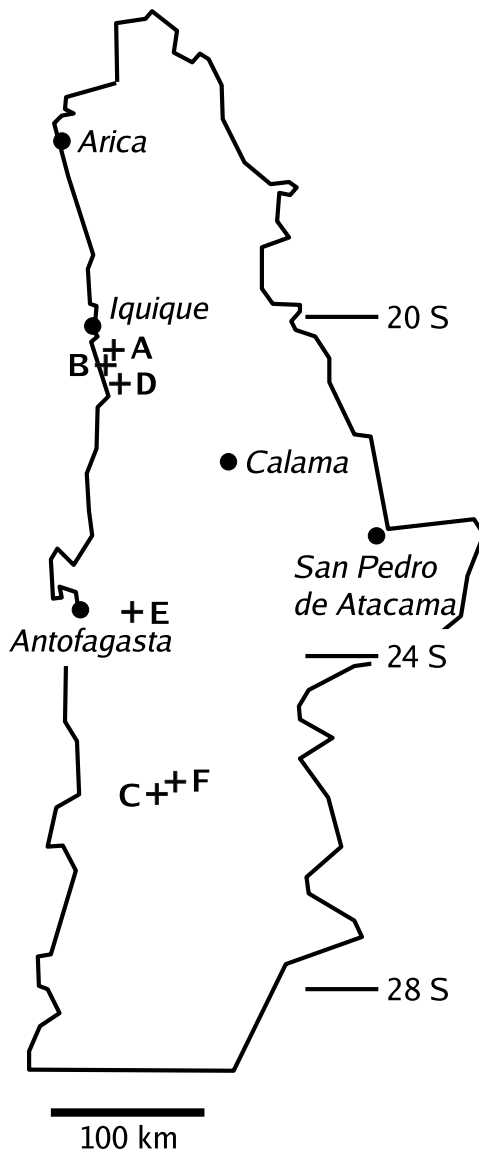
## 2. Robotic Investigation

[5] Beginning in 2005 Carnegie Mellon’s “Science on the Fly” project investigated science autonomy technologies for use in planetary science applications. It focused on capabilities for science operations that involve long over-the-horizon traverses, including autonomous detection and classification of rocks, autonomous spectroscopy, science goal representations and planning. A partnership with the LITA field campaign provided a unique opportunity to test autonomous science technologies as an integrated part of an exploration robotics field campaign.

<sup>1</sup>Robotics Institute, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA.

<sup>2</sup>NASA Ames Research Center, Moffett Field, California, USA.

<sup>3</sup>Molecular Biosensor and Imaging Center, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA.



**Figure 1.** Map showing the location of LITA field campaign sites [from Cabrol *et al.*, 2007].

[6] The LITA project aimed to survey the distribution of extremophile life and habitats in the Atacama, a Mars-analog desert in Chile. The three Atacama field campaigns visited six field sites (Figure 1): sites A, B and D were located in the “humid zone” near the Coastal Range; sites C and F were in the arid core of the desert; site E was deeper into the central depression. Scientists explored the desert remotely using an autonomous rover commanded from North America. Remote operations enforced space-relevant constraints such as limited bandwidth and a single daily command cycle. This permitted research into extremophile habitats while simultaneously testing operational strategies to exploit navigational autonomy in fast rovers. For a more complete overview, refer to Cabrol *et al.* [2007].

[7] The robotic exploration utilized Zoë, a solar-powered rover capable of single-command autonomous navigation of multiple kilometers (Figure 2). Its instrument suite included

a mast-mounted pan-tilt unit with narrow-field panoramic cameras (pancams) and fixed forward-facing navigation cameras (navcams). It also carried an integrated microscopic imager capable of detecting chlorophyll or bacterial colonies. An onboard sprayer deposited fluorescent dyes which would bond to living organisms after a short incubation period (Figure 3). The microscopic imaging apparatus would induce fluorescence with a flashlamp. Subtracting ambient light left only the fluorescence from bonded dyes, an in situ image of any bacterial colonies that were present. Finally, the science team used orbital imagery from ASTER [Abrams, 2000], Hyperion [Pearlman *et al.*, 2001], and IKONOS [Dial *et al.*, 2003] as an integral part of planning over-the-horizon traverses. Details about science instrument payload are given by Cabrol *et al.* [2007] and Weinstein *et al.* [2007].

[8] The Atacama campaign provided an appropriate test bed for science autonomy technology. These technologies were integrated into rover operations and evaluated as a part of a broader rover exploration strategy.

[9] The science goals of the LITA project were particularly amenable to science autonomy. The scientists attempted to complete a wide-area biogeologic survey in a short time by visiting new environmental units whenever possible. The data collected during these traverses involved areas that scientists had not seen previously. However, they expected that extremophile habitats would be distributed in isolated patches and oases of life as inferred from a “follow-the-water” strategy. Thus the data collected during a long traverse contained a few positive images containing life mixed with a large number of negative images. The extensive measurements required to verify life at any single location were expensive in terms of rover time and limited amounts of experimental dye resources. It was hypothesized that science autonomy could manage these resources more efficiently by performing life-detection experiments selectively in response to more expedient observations.

[10] Finally, the Atacama desert environment permitted tests of site characterization. Like Mars, the Atacama is bare of vegetation; site properties such as the distribution of surface materials were plainly visible to rover cameras. Moreover, the long traverses of the Atacama project meant that Zoë often traveled across multiple geologic and environmental units. These data sets provide a unique opportunity to test automatic site profiling strategies in a Mars-analog environment.

[11] We present experiments in four specific autonomous science domains that illustrate a range of science autonomy options available to rover designers. Section 3 demonstrates autonomously responding to evidence of life with more detailed measurements, specifically reacting to chlorophyll fluorescence detected using a microscopic imager. Section 4 considers rock detection for site profiling and selective data return, using the visible properties of rocks as a support to geological and morphological characterization of the sites. Later sections deal with science phenomena on larger scales (Figure 4). Section 5 details tactical replanning to map the distribution of life. Here an active mapping system actually moves the rover to new locations in response to collected data. Finally section 6 investigates detection of novel images and geologic unit boundaries in image sequences

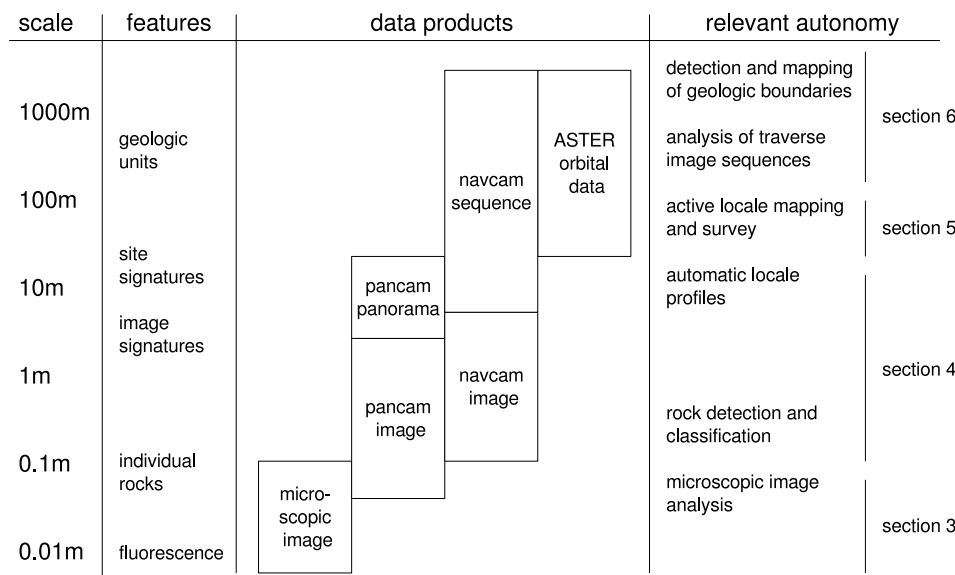


**Figure 2.** The Zoë rover platform in the Atacama desert.



**Figure 3.** The microscopic fluorescence imager deployed and spraying underneath the robot.





**Figure 4.** Science autonomy experiments in the LITA project dealt with instruments and features at multiple scales.

to support selective data return. In each domain area quantitative comparisons demonstrate that science-aware decision making leads to improvements in the quality of returned rover data.

### 3. Autonomously Responding to Evidence of Life

[12] The LITA campaign strategy for unambiguously detecting life was to search for spatially correlated fluorescence signals from several distinct types of biomolecules using the robot's onboard microscopic fluorescence imager (FI). The full protocol of FI measurements at a locale was resource-intensive, taking approximately 23 minutes and using fluorescent dye from a limited onboard stock.

[13] However, since the emphasis was on finding locales with positive signal from as many distinct biomolecules as possible, we were able to improve overall efficiency by terminating the protocol early if initial signals came back negative. In this way we saved rover resources without sacrificing high-value unambiguous detections of life.

[14] We divided the full protocol into two segments: (1) Check quickly for chlorophyll; (2) Only if chlorophyll was detected through onboard image analysis, autonomously respond by performing followup measurements with dyes to detect other biomarkers. In the common case that the result from step (1) was negative, the modified protocol took only eight minutes to complete and did not consume any dye. Our field evaluation confirmed that the improved protocol significantly improved overall efficiency.

#### 3.1. Fluorescence Imager (FI) Instrument

[15] The FI is a down-pointing camera mounted on the bottom of the Zoë rover (Figure 3). It has 10 cm field of view and transverse resolution 210  $\mu\text{m}$ . During autonomous response experiments, the sampling location under the FI was chosen by stopping the rover at fixed distances along its

traverse, and the camera was deployed and auto-focused using z-axis motion.

[16] The FI could detect either the reflectance or fluorescence of a sample in various channels. A xenon flashlamp provided illumination. Servos could select one of six optical interference filters for the excitation path between the flashlamp and sample, and one of ten filters for the detection path between the sample and CCD.

[17] The FI captured reflectance under a combination of sunlight and flashlamp illumination with no excitation filter. In RGB color mode, separate images with red, green, and blue emission filters combined to create a visual color image. In fluorescence mode the FI captured a greyscale intensity image with excitation and emission channel pair selected to respond to the fluorescence of the molecule under study (either chlorophyll or an artificial marker dye which fluoresced when bound with a biomolecule). Different marker dyes responded to DNA, proteins, lipids, and carbohydrates. An automatic sprayer could spray the sample with water to enhance chlorophyll fluorescence under dry conditions. It could also spray a solution that contained all four marker dyes along with acid and detergent to aid dye penetration. For more details on the FI and its effectiveness in the field, see *Weinstein et al.* [2007] and *Warren-Rhodes et al.* [2007].

#### 3.2. Chlorophyll Detection Experimental Procedure

[18] The FI protocol for each sample had two phases: (1) The FI sprayed water and captured several images, including a chlorophyll image used to determine if followup was warranted. (2) It sprayed the dye mixture and captured several images, including dye fluorescence images. Execution times for the phases were approximately 8 and 15 minutes, respectively.

[19] Scientists could command two types of sample (Figure 5). A full sample always included both phases of the protocol. A periodic sample always included phase 1,

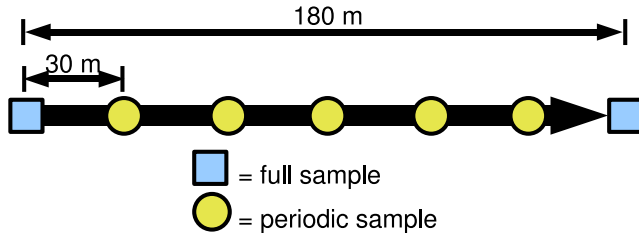


Figure 5. Periodic sampling traverse.

but continued to phase 2 only if chlorophyll was autonomously detected (Figure 6).

[20] Scientists used the autonomous followup capability during 180 m periodic sampling transects (Figure 5). At each endpoint of the transect, the rover stopped and took a full sample. At 30 m intervals within the transect, the rover stopped and took a periodic sample. The rover executed each traverse autonomously within a single command cycle (including both driving and FI samples).

[21] In keeping with the LITA high-mobility exploration strategy, the science team decided to add a followup quota ensuring that the rover would not dwell too long on a particular transect. At most three periodic samples per transect could trigger followups. After filling this quota the rover would not spray dye on subsequent samples even if it detected chlorophyll.

### 3.3. Chlorophyll Detection Image Analysis

[22] Autonomous chlorophyll detection, used only for periodic samples, relied on a single image of chlorophyll fluorescence intensity (excitation 450 nm, emission 740 nm) captured after the FI sprayed water on the sample. This was called the “trigger image” (Figure 7).

[23] The detection algorithm reported the probability that chlorophyll was present anywhere in the image, triggering an autonomous dye followup if the probability was 50% or higher. The algorithm reported a high probability if there were any bright patches in the image. First it formed a subsampled image by splitting the original image into  $4 \times 4$  cells and calculating the mean intensity  $A(x, y)$  over each cell. This smoothing eliminated false detections from single-pixel shot noise. Second, it converted mean intensity for each cell to a probability  $L(x, y)$  that the cell contained chlorophyll using a sigmoid or “fuzzy threshold” function  $\sigma_{\alpha\beta}$ . Finally, it calculated the overall probability  $L$  of chlorophyll being present in any cell of the image by combining the probabilities from individual cells using the heuristic assumption that the  $L(x, y)$  measurements were uncorrelated.

[24] Let  $I(x, y)$  denote the pixel intensity at position  $(x, y)$  in the original image. We have

$$A(x, y) = \frac{1}{16} \sum_{\Delta x=0}^3 \sum_{\Delta y=0}^3 I(4x + \Delta x, 4y + \Delta y) \quad (1)$$

$$L(x, y) = \sigma_{\alpha\beta}(A(x, y)) = (1 + \exp(\alpha + \beta A(x, y)))^{-1} \quad (2)$$

$$L = 1 - \prod_{x,y} (1 - L(x, y)). \quad (3)$$

[25] The fuzzy threshold function  $\sigma_{\alpha\beta}$  used to convert cell intensity to probability of containing chlorophyll had two parameters  $\alpha$  and  $\beta$  whose values were tuned using training data. The training data were gathered by manually labeling the presence or absence of chlorophyll signal in individual  $4 \times 4$  cells of two trigger images containing lichens, based on morphological cues from both the trigger image and an associated visual color image.

[26] Let  $n = 2$  denote the number of training images, let  $A_i(x, y)$  denote the mean intensity in cell  $(x, y)$  of the subsampled version of the  $i$ th labeled image, and let  $L_i^*(x, y)$  denote the corresponding manual label, with value 1 or 0 indicating the presence or absence of chlorophyll in the cell. The parameters  $\alpha$  and  $\beta$  were set to maximize the likelihood of the manual labels using logistic regression.

$$(\alpha, \beta) = \arg \max_{(\alpha', \beta') \in \mathbb{R}^2} \prod_{i=1}^n \prod_{x,y} (1 - |L_i^*(x, y) - \sigma_{\alpha'\beta'}(A_i(x, y))|) \quad (4)$$

[27] While every labeled training image contained significant instances of chlorophyll fluorescence, many images from the Atacama Desert did not. To account for this discrepancy we included an additional parameter  $\tau$  that represented a prior on the probability of finding chlorophyll in the image (equivalently, the proportion of the ground surface expected to be covered by photosynthetic organisms). This transformation, applied after training is complete, yields

$$L(x, y) = \frac{\tau \sigma_{\alpha\beta}(A(x, y))}{\tau \sigma_{\alpha\beta}(A(x, y)) + (1 - \sigma_{\alpha\beta}(A(x, y)))} \quad (5)$$

[28] The value of  $L$  was then calculated from  $L(x, y)$  using equation (3) as before. We informally hand-tuned  $\tau$  on a set of test images so that most images would fall on the correct side of the followup threshold (negative images below 50%,

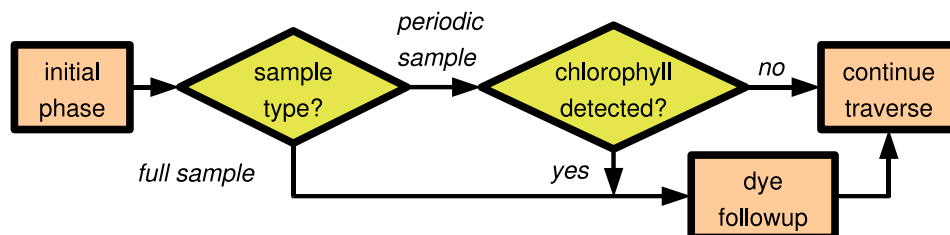
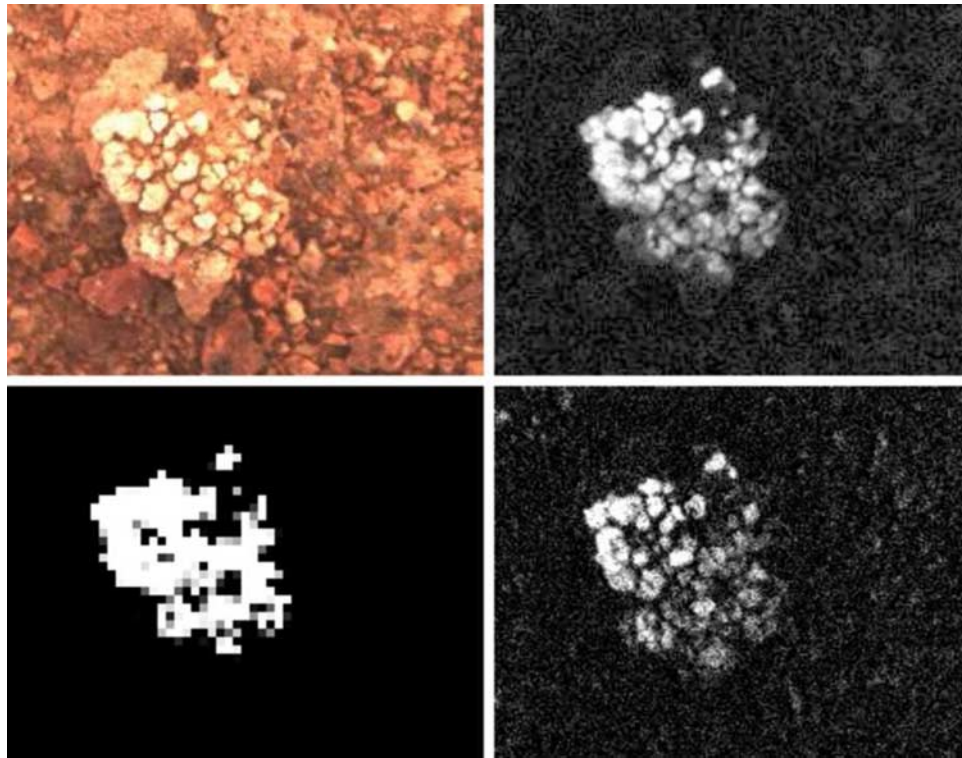


Figure 6. Sample protocol flowchart.



**Figure 7.** (top left) Portion of FI visual color image containing a lichen. (top right) Chlorophyll trigger image. (bottom left) An intermediate step of image processing; the brightness in each cell represents the estimated probability that it contains chlorophyll. (bottom right) After autonomous followup, the FI detected fluorescence from the DNA marker dye.

positive images above 50%). The resulting value of 0.005 was used for all of the reported results.

### 3.4. Chlorophyll Detection Experimental Results

[29] The autonomous response system was evaluated onboard the rover during the 2005 Atacama field campaign. Our data set included 24 periodic samples collected during five traverses. Since the standard traverse length was 180 m, and periodic samples occurred at 30 m intervals, nominally each traverse should have included five periodic samples. In practice, the number of samples varied slightly because of small positioning errors in the commanded endpoint locations from the science planning interface. Nightfall ended the last traverse after only two periodic samples. In two of the traverses the rover filled the followup quota before the last periodic sample; we excluded later periodic samples from the analysis because the quota prevented any further followups.

[30] Each sample image set was analyzed by a remote team that included field biologists and fluorescence experts. Using both the visual color image and the trigger image, they labeled the samples as positive (contains significant evidence of life) or negative (does not). We compared the scientist labels to the autonomous followup response.

[31] The autonomous system and expert labels agreed for 19 of the 24 samples. 8 of the 24 samples were positive; 7 of the 11 samples chosen for autonomous followup were positive. Thus the yield, or proportion of dye samples that occurred after positive evidence of chlorophyll, was 90% higher when applying the science-aware system than would

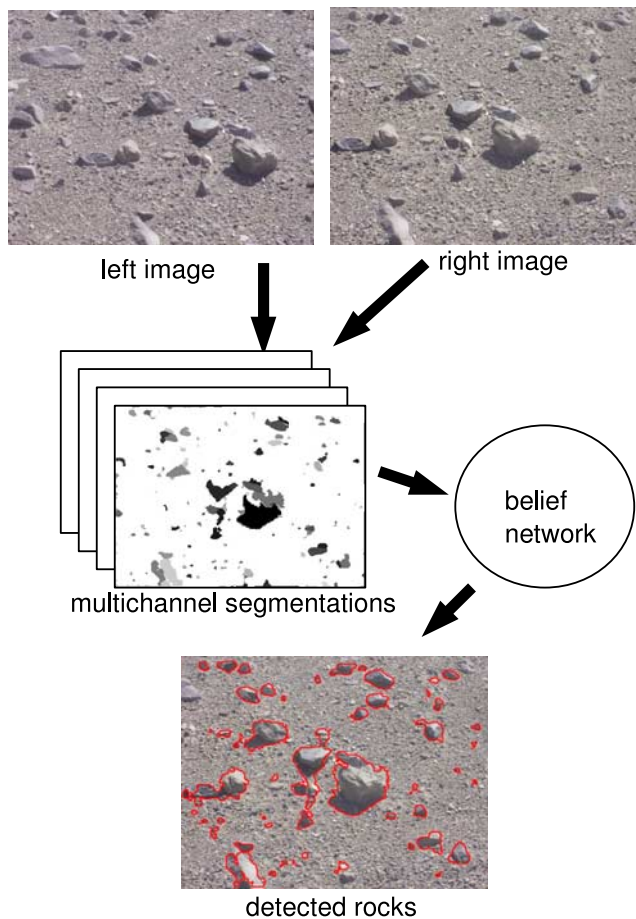
have been expected for randomly selected samples (significance level  $< 0.01$  using one-tailed Fisher's exact test). In other words, relative to a baseline strategy of always applying dye the science-aware system skipped 13 dye samples, saving more than three hours of robot time, while following up chlorophyll detection correctly in 7 out of 8 cases.

[32] Failure modes we observed over the course of the field campaign included corrupted chlorophyll signal when the FI was dazzled by direct sunlight (normally it is shaded by the robot), and failure of the water spray to reach the sample due to high winds. These failures could be addressed by hardware modifications or by limiting FI operations to certain times of day. The image analysis could also be made more robust to corrupted images by diagnosing problems like dazzling and poor focus based on the image, or by adding more definitive cues such as squamous lichen morphology to the analysis.

## 4. Rock Detection for Site Profiling and Selective Data Return

[33] One can also apply adaptive collection and return principles to macroscopic geologic and biologic features. During the LITA project we conducted several experiments concerning the identification of rocks in Atacama Desert imagery. Rocks are particularly valuable targets that contain useful geologic and potentially biologic information (e.g., endoliths, chasmoliths) across a wide range of sensing modes. They are important candidates for autonomous





**Figure 8.** Procedure for rock detection.

instrument deployment and selective data return. This makes them a natural test case for geologic and biologic feature detection. LITA's need for fast, robust rock detection in a broad range of environments led us to develop a new detection algorithm.

[34] Testing our system with Atacama data, we found that rocks over 20 pixels in size could be found with up to 70% precision at 60% recall. We also demonstrated the use of rock detection for computing numerical site signatures in order to compare neighboring locales. In a 200 m test traverse, automatic rock detection and classification yielded site signatures that correlated well with different distributions of surface material observed at the site. Finally, we investigated the use of detected rocks in guiding selective data return decisions. A blind survey of LITA project scientists evaluated their preferences for images selected automatically against a blind sampling strategy. This survey indicated a significant preference for intelligent data return utilizing rock detection.

[35] Several past systems have used stereo images to find rocks based on their height above the ground plane [Gor *et al.*, 2000; Fox *et al.*, 2002; Pedersen, 2002]. These methods can reliably find large rocks on flat ground, but they cannot usually identify small or partially buried rocks and their range is limited to around 10 m. A second method reduces rock detection to the simpler problem of finding shadows [Gulick *et al.*, 2001]. This strategy finds a point on the

rock's surface but does not find the outline of the rock, which can make subsequent classification more difficult. Several other methods detect rocks based on their contrast with the background in terms of albedo, color, or texture. For example, Castaño *et al.* [2004] look for closed shapes with an edge detector. This technique works best for finding a small number of rocks with high accuracy. Our new algorithm borrowed from these previous systems and used any available intensity, shading and height information to help inform detection decisions.

#### 4.1. Detection Algorithm

[36] Given the wide variety of rocks, background soil, and lighting conditions in the Atacama, we wanted a rock detection system that could incorporate multiple types of visual cues. This led us to develop a new rock detection algorithm consisting of two steps: "feature segmentation" which segmented images into subregions of constant properties, and a "detection" stage which used any available sensor data to identify rock regions from among the candidates.

[37] The segmentation step processed sensor data to identify homogeneous regions in the rover's environment corresponding to possible rocks. The segmentation stage was not responsible for detecting targets but merely suggested candidate regions for further classification. This separation between segmentation and detection permitted the rock detection system to be quite permissive in the types of data it accepted as input. Any additional criteria, even from multiple measurement sources, could add candidate regions for consideration.

[38] The segmentation procedure employs a simple region-merging strategy. It shatters the image into a grid of  $5 \times 5$ -pixel squares and iteratively joins them back into regions of uniform properties. At each iteration we calculate the mean pixel values of all regions; contiguous regions whose means fall within a certain threshold of each other are merged. This process repeats until no more merges occur. This entire procedure operates separately on hue, saturation, and intensity color channels separately to yield multiple independent segmentations of the same scene (Figure 8). A feature which was indistinguishable from the background in one channel was sometimes visible in another; a rock whose contour appears in at least one segmentation still had a chance to be recognized later in the detection stage.

[39] For some tests stereo imaging provided additional geometric data in the form of a 3D point cloud. In these cases we converted the 3D data into a height map which constituted an additional fourth channel for the segmentation procedure. We computed the height map by fitting a squared-error-optimal plane to the point cloud and then calculating the physical distance from each pixel to the groundplane. Applying the region-merging algorithm to the height map produced additional candidate rocks of uniform height. Our simple planar terrain model worked best for large rocks and flat ground; any undulations in the terrain would dominate the height segmentation and nullify the benefits of stereo imagery.

[40] The detection stage identified science targets from among the segmented regions by extracting a real-valued attribute vector from each candidate and labeling it with a Bayesian belief network [Pearl, 1988]. The belief network

is just one of a wide variety of supervised learning techniques that one can train with exemplars to label new data points. For our application belief networks offered some advantages over comparable techniques. First, they solved the problem of missing data—for instance, in the case of a stereo mismatch, the network provided posterior probabilities that were appropriate for the available information. Another advantage to the Bayesian approach is its transparency; unlike many other supervised labeling methods it is easy to diagnose and explain the reasons for each decision.

[41] During the detection stage the belief network was labeled using a vector of numerical attributes:

[42] 1. Perimeter: The ratio of the region's squared perimeter to its pixel area. Nonrock artifacts often had long spidery shapes while rocks were more convex and ellipsoidal.

[43] 2. Relative color: The absolute value of the difference in mean pixel hue, saturation, and intensity between the interior of the region and its local context as defined by an enclosing rectangle.

[44] 3. Relative color variance: The absolute value of the difference between the pixel variance of the interior region and the pixel variance of its local context.

[45] 4. Height above the ground plane: While small rocks are generally below the noise threshold for our stereo system, height is a valuable attribute for detecting large rocks and excluding large nonrock regions.

[46] 5. Texture: We use a fractal dimension measure [Chaudhuri and Sarkar, 1995] of a binary intensity map to describe the detail of each region as resolution is increased. The result is an efficiently computed value that corresponds somewhat to our intuitive notion of surface roughness. It is worth noting that Dunlop [2006] has recently shown texture analysis to outperform fractal dimension for this task.

[47] 6. Intensity gradient: Rocks are three-dimensional protrusions that exhibit shading when illuminated by sunlight. We used least-squares regression to find the magnitude of the intensity gradient over the pixel surface of each region in the image, giving the overall strength of its shading.

[48] Two additional attributes do not directly affect a region's chances of being a science target or its geological classification. Nevertheless, we included them due to their strong conditional dependence relationships with the other attributes. By considering these dependencies the Bayes network computed a more accurate posterior probability:

[49] 1. Absolute range: Many of the differences between rocks and nonrocks become less apparent as range increased. In particular, texture was more difficult to see even when there were fewer pixels representing the distant region.

[50] 2. Pixel area: There were varying degrees of conditional dependence between most observed attributes and the regions' pixel area. These dependencies were due to the way regions are represented as a finite number of pixel "samples." For example, normalized perimeter was generally small for regions of small pixel area because rough borders became less noticeable when the number of pixels used to describe them decreased. Similarly, texture was hard to recognize with few pixel samples.

[51] The belief network used any available values from this attribute vector to calculate the probability of a region belonging to each of five classes: rocks, uniform patches of soil, sky, shadows cast on the ground, and a final "everything else" class for regions that did not fit neatly into one of the other categories. The "everything else" class contained ambiguous candidates like a region comprising a small part of a larger rock or one that included both a rock and a small patch of sediment. We estimated belief network parameters using examples of region classes based on manually labeled segmentations of training images. Conditional probabilities were calculated based on the empirical probabilities exhibited in this training data. At runtime the belief network attempted to replicate the class labeling for novel data, resulting in a list of image regions corresponding to likely rocks. For a more complete explanation of the rock detection system, see Thompson *et al.* [2005a].

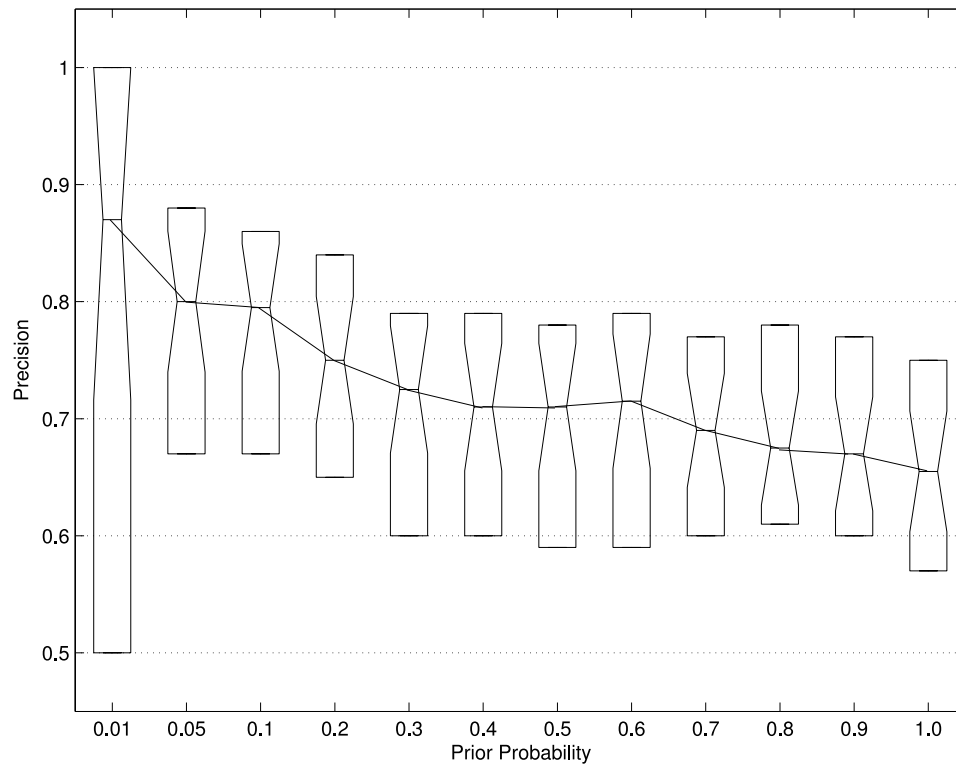
#### 4.2. Rock Detection Accuracy: Experimental Procedure and Results

[52] We tested the rock detector's performance on a pair of panoramas taken from rover imagery in the Atacama Desert. We chose a hill at site C strewn with rocks of various sizes to serve as a test site. Data collection occurred at two locales: a training set captured at the base of the hill and a test set captured at its peak. At each site the rover used its stereo imaging suite to capture partial panoramas comprising 180 degrees of azimuth and 40 degrees of elevation. We manually labeled segmentation outputs from the first panorama used them to train the region detection stage. Then the fully trained detection algorithm identified rocks in the second panorama.

[53] We evaluated these autonomous detections against the locations of real rocks in the image set, restricting our definition of a rock to those that had a short axis longer than 20 pixels. To calculate a performance score we checked that the center of each detected bounding box lay within the true contour of the rock as drawn by hand. If the center fell within an unclaimed true rock contour the detection was a true positive. This matching criterion does not consider the degree to which the detected rock's border contours match the actual rock, which is a much harder problem. We scored each image by matching true rocks in the scene against available detections until no more matches were possible. Any remaining true rocks were false negatives, and unasociated detections were false positives.

[54] Figures 9 and 10 illustrate the detector's performance on the 1078 coded rocks in the test data. Varying the prior probability for the "rock" class results in different performance characteristics. Precision, the fraction of detected targets in coded regions that correspond to actual rocks, is in general inversely related to recall, the fraction of actual targets that are found. Figure 9 illustrates detection precision for varying choices of the class prior. The notches in the box plot represent the 95% confidence interval for the median, treating each subimage of the panorama mosaic as an independent sample. As the prior probability increases a greater number of uncertain regions are classed as science targets and precision drops. Figure 10 shows the recall rate for each prior probability. At a prior probability of 0.5, 60% of all true rocks are found and 70% of all detected rocks



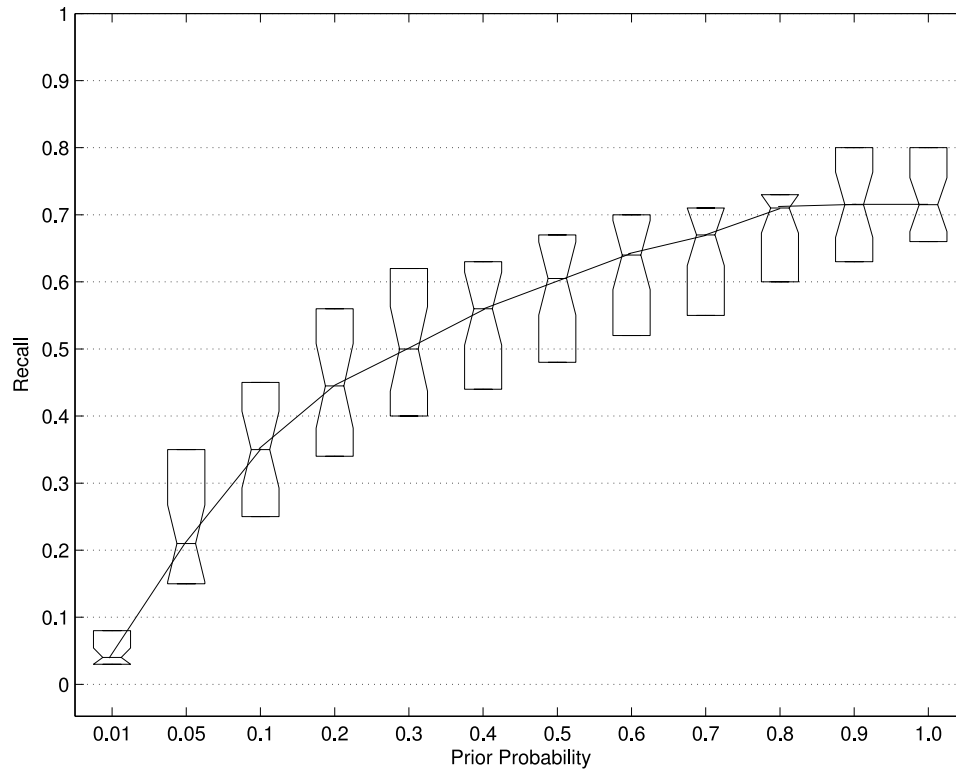


**Figure 9.** Rock detection precision on the test panorama.

(according to the correspondence metric outlined above) are correct.

[55] Note that the detector never found all of the science targets, even for the case where the prior was set to 1. The

reason is that our evaluation measures not simply the accuracy of the region “rock” or “nonrock” labeling but rather the overall system’s ability to find rocks in the original images. A rock that does not appear in the seg-



**Figure 10.** Rock detection recall on the test panorama.

mentation stage has no chance of being labeled later. The rock detection results of this test do not necessarily mean that this rock detection system would achieve the same results on any terrain. The training set was taken from images captured at the same general area as the test set. Thus the experiment does not prove the detector would perform as well on very different rocks or terrain. However, scientists would always be able to retrain a detector using appropriate training data from any new region the rover entered.

[56] LITA's long-distance driving strategy provided a stiff challenge for the rock detection system, requiring it to accurately detect rocks in natural environments and under field lighting conditions. These experimental results demonstrate the effectiveness of a principled machine learning approach that can flexibly integrate information from a variety of visual cues and multiple sensors. Moreover, given that the Atacama is a Mars-analog environment and the detector can be quickly adapted to new data sources using training examples, there is reason to believe the same techniques will apply to Mars rover imagery.

#### 4.3. Relating Detected Features to Scientist Priorities

[57] In the most arid regions of the Atacama extremophile life is sparse and limited to benign microhabitats such as crevices within or beneath translucent rocks [Friedmann and Galun, 1974; Warren-Rhodes *et al.*, 2006]. In order for the rover to realize that a potential microhabitat is worthy of further study we need some way for the science team to express their priorities in terms of features the rover can understand. Our approach to this problem relies on concepts from previous work.

[58] Chien *et al.* [2005] describes a simple preference model that does not distinguish the relative priorities of different feature classes. This strategy was used for the Autonomous Sciencecraft Experiment conducted aboard the Earth Observing One orbiter. Transient phenomena such as floods, fires, and volcanic activity were captured using autonomous retargeting of a high-resolution imager, where manual retargeting would have been too slow.

[59] Castaño *et al.* [2003] proposed three more expressive models for science value. Scientists using the target signatures strategy assign priority scores to particular classes of features. If the rover downlink budget is oversubscribed, the rover can use the scores to return the most informative subset of the collected data. This approach has the disadvantage that scientists can only prioritize classes of features that they can anticipate in advance. Using the representative sampling strategy, the rover builds a statistical model for observed features on the fly. It groups the data into clusters with the object of returning a data set containing examples from each significant group of observed features. Using the novelty detection strategy, the rover prioritizes features that do not fall into any known class, presuming that these features are interesting precisely because they are unusual.

[60] Smith *et al.* [2005] proposed an integrated approach encompassing all three of these strategies. It worked by supplementing prespecified target signatures with new target signatures for unanticipated classes. The new target signatures carried priorities based on scientist interest in representative sampling and novelty detection.

[61] We also note that the artificial intelligence community has devoted considerable study to the general problem of eliciting and representing complex preferences from human decision-makers; see Chen and Pu [2004] for a review. Their techniques deserve more attention from the science autonomy community.

#### 4.4. Geologic Classification

[62] The study of rocks was central to the LITA project's objectives. Rocks' size and shape indicate geological processes and environmental transitions. Rocks themselves can be hosts to life. However, rocks are numerous and the science team had only limited operation time. As a result there was a clear need for some tasks to be made autonomous. In order to assist the science team in the characterization of the geological environments as potential habitats to life, we developed an automatic rock classification system for LITA that could categorize rocks using visual attributes. After finding rocks in all the images of a traverse the system used the numerical attribute vector described in section 4.1 to sort detected rocks into geologic classes. To offer scientists more flexibility we permitted categories according to both supervised and unsupervised definitions as in Smith *et al.* [2005].

[63] Our study built on the growing body of research demonstrating automatic classification of rocks. Given rock outlines, one can automatically calculate shape measures like circular variance and Diepenbrök roundness. Dunlop [2006] suggests that these measures correlate well with Crofts' roundness and sphericity measurements, which in turn are useful indicators of how rocks were created and transported. Spectral properties also offer mineralogical information for geologic classification. The Robotic Antarctic Meteorite Search of Pedersen *et al.* [2001] integrated spectroscopy and visual imagery to find and classify meteorites during robotic traverse on the Elephant Moraine plateau. Pedersen [2001] demonstrates statistical models of the local environment that use contextual cues to improve classification performance by the explorer robot. More recently, Bornstein and Castaño [2005] have demonstrated automatic classification of carbonate minerals using spectral characteristics.

[64] Our classification system permitted several methods of defining geologic classes:

[65] 1. Interval classes: These classes specified a binary decision boundary along specific attribute values chosen by the scientist. For example, quartz rocks host microbial communities due to their translucence and can sometimes be identified by their high albedo; a scientist could instruct the system to label as quartz any rock which had an albedo greater than two standard deviations above the mean. Only rocks that did not fall into a user-defined interval became candidates for one of the other classifications.

[66] 2. Example classes: We modeled this class with a multivariate Gaussian distribution over the vector of rock attributes. Scientists could specify an example-based class by providing one or more examples. A maximum-likelihood strategy [Bishop, 1995] then provided the best parameters of the distribution describing these examples.

[67] 3. Automatic classes: These classifications were similar to example classes in that they are defined by a multivariate Gaussian. However, these clusters were gener-

**Table 1.** Number of Times a Preference Was Expressed for Either Sampling Strategy

Data Return Method	Panoramas	Short Sequences	Long Sequences
Periodic	9	23	12
Selective	15	49	26

ated automatically in response to new data. An expectation-maximization (EM) algorithm [Dempster *et al.*, 1977] fit automatic cluster parameters to the entire data set while leaving example class parameters fixed. In this manner the automatic classes shifted to accommodate data that was inadequately described by other classes.

[68] 4. Outliers: We modeled novel rocks with a uniform probability density function that collected examples not belonging to one of the other classes. By altering this density scientists could be more or less permissive about what counted as an outlier.

[69] These options provided an expressive vocabulary of possibilities for defining geologic classes. Nevertheless, the simplicity of the visual properties available (color, texture, etc.) meant that these classes could probably never correspond perfectly to formal geologic or mineralogic classifications. We hypothesized that even without perfect geologic fidelity, a loose correlation of automatic classifications with the real categories of interest would improve data collection over random or periodic measurement techniques.

#### 4.5. Selective Data Return: Experimental Procedure and Results

[70] For LITA—as in any future rover mission dealing with large quantities of collected data—data return prioritization was a critical issue. We conducted a blind survey of 8 LITA project scientists to evaluate their preferences for or against image sets selected using intelligent sampling techniques. We processed several data sets using both a selective return strategy and periodic measurement strategies that ignored image content. The selective return strategy detected rocks in all images and classified the detected rocks. It then returned the best images as identified by a selective return algorithm.

[71] We chose several site C data sets for their preponderance of rocks and diversity of terrain. The scientists evaluating the data sets for this experiment were kept blind: this was their first view of the locale in question, although they had previously viewed imagery from nearby areas with similar geology. The data sets were chosen to simulate real data return tasks that might occur in field work:

[72] 1. Four 270-degree panoramas that excluded the rear quadrant. These panoramas each contained 85 images. The scientists were given a reduced-resolution thumbnail of the entire panorama; the data return task consisted of finding the best three high-resolution images to transmit. The baseline periodic method returned the forward, right, and left-looking images at a 30-degree declination. This strategy appeared to provide the best compromise between detail and context.

[73] 2. Six long traverse segments of 100 pancam images each, sampled at 2m intervals. These six segments were

taken from two long drive actions near site C. The pancam faced forward at a 15-degree declination for all images. No thumbnails were provided. The task consisted of returning the best image frames from the sequence. The baseline strategy selected the first, last, and middle frames from each traverse.

[74] 3. Ten short traverse segments of 30 pancam images each, sampled at 2 m intervals. Again, no thumbnail images were provided. These segments came from a single long drive action with similar pancam imaging parameters. The baseline strategy selected the first, last, and middle frames from each segment.

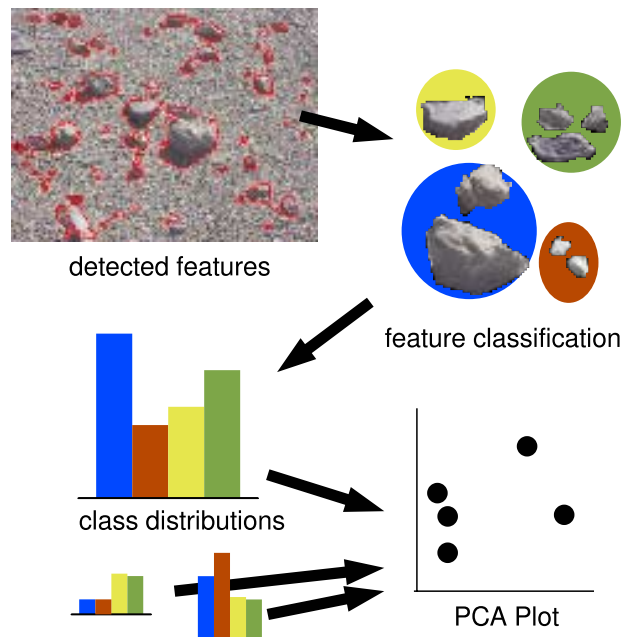
[75] The LITA intelligent return system used a combination of both representative sampling and novelty detection. For each data set it first computed 10 automatic classes using standard clustering techniques. It then assigned a score of 5 to each image containing at least one outlier rock and a score of 1 for each automatic class represented in an image. Thus with 10 automatic classes the maximum possible score for a single image was 15. However, a “neighbor threshold” distance of 5 m prevented any two nearby images from getting scores for examples of the same class. In theory, this would prevent any single feature from monopolizing the downlink. In addition, a “quantity threshold” indicated the maximum percentage of the downlink that could receive a score for any particular class. This threshold was set to 0.25 for the outlier class and 0.1 for each automatic class. The quantity threshold also helped to encourage a diversity of different images. These thresholds and the parameters of the EM algorithm were tuned manually to achieve reasonable results on a similar set of data to the test set described above. A greedy data return algorithm chose the three downlink images one at a time, each time adding the top-scoring image, in order to maximize the overall value of the data set.

[76] For each of the 20 data sets the scientist was given two triplets of downlinked images in random order. One triplet had been chosen through periodic data return, while the other was chosen through automatic selective return. Occasionally both methods chose the same image by coincidence; in these cases the repeated image appeared in both triplets. The scientist was asked to choose the downlink triplet that “they would rather receive to help them understand the traverse area.” Scientists responded by expressing a preference for the first triplet, the second triplet, or neither.

[77] The complete survey consisted of 160 trials. Table 1 summarizes the number of times any preference was expressed for either return method. The total number of observed preferences was 134, out of which 90 favored intelligent sampling and 44 favored periodic sampling (significance level  $< 0.001$ ,  $\chi^2 = 15.79$ , against the hypothesis that scientists would choose either method with equal likelihood).

[78] While scientists showed a significant overall preference for selective data return they usually provided reasons for their choice that had little to do with the classes of rocks in the scene. These reasons varied broadly depending on the scientist’s own background and the perceived task at hand. Occasionally a scientist would prefer a high-angle panorama tile because it provided more context for planning navigation. Other times scientists focused on interesting patches of





**Figure 11.** Procedure for automatic site profiling.

soil or a curious shadow that indicated a larger object outside the image frame.

[79] Scientists' overall preference for selective return is interesting given the relative unimportance of rock classifications in their stated rationales (which of course will vary according to the perspective of each mission). It is likely that the rocks chosen for selective return were correlates of these other interesting phenomena. The selective return algorithm favored images with lots of rocks. Moreover, the representative sampling component favored a broader diversity of image content. These factors filtered out homogeneous stretches of empty terrain that often dominated the periodic data sets. In the case of panoramas, the selective return strategy may have inadvertently gained an advantage with the potential to return both near-field and far-field imagery. This often led to a greater diversity of elevations in the downlink. In many cases the scientist's decision criteria was not addressed explicitly by either selection method. Nevertheless, the overall preference in favor of intelligent return suggests that rock detection and classification improved the quality of the intelligently selected data sets.

#### 4.6. Rock Detection and Classification for Autonomous Site Profiling

[80] A natural extension of autonomous rock detection and classification was the generation of automatic geologic signatures—profiles describing the distribution of rock classes at a site [Thompson *et al.*, 2005b]. These geologic signatures reveal subtle geologic trends, the borders between geologic units, and sudden changes compared to neighboring locales. All of these onboard analyses can inform adaptive measurement or return decisions. An experiment analyzed a traverse consisting of five neighboring locales at LITA site C. It suggested that site profiles based on autonomous rock detection correlated with the principal distinctions in surface composition identified by a human at the site.

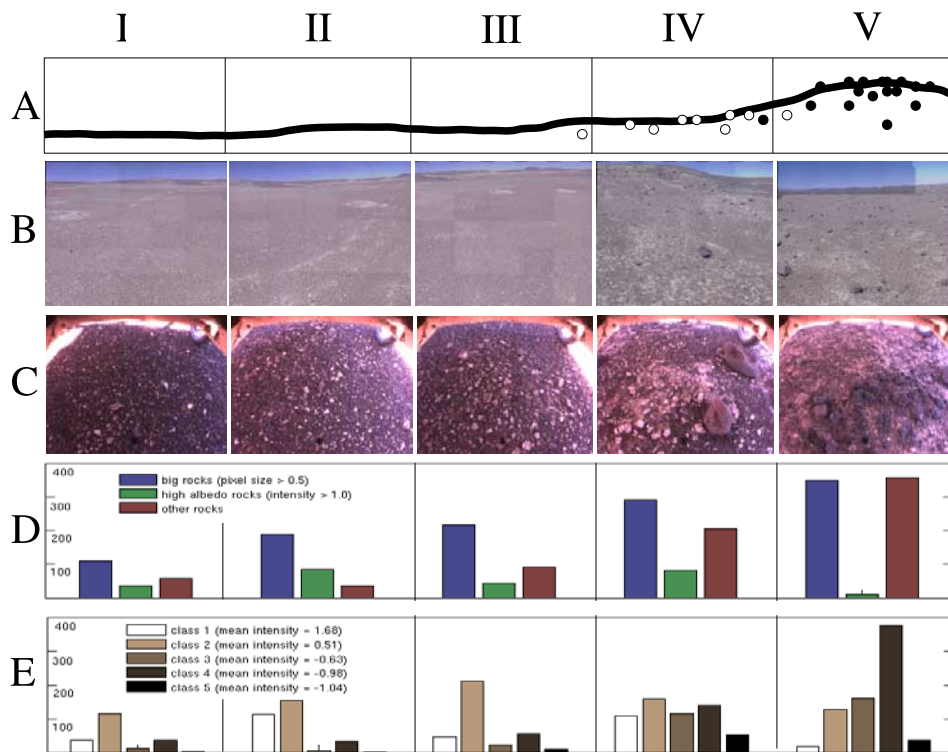
[81] The site profiling procedure consisted of several stages (Figure 11). The technique first detected rocks in all images and then classified them according to automatic or user-defined classes. The result was the quantity of each different class at a locale. Subtle changes like a shift in the density of a certain class identify unit borders and suggest areas for further exploration. For visualization we applied principal component analysis (PCA) to the class distribution histograms and projected them onto their first two principal components. This provided a compact 2D visual comparison of the different locales.

[82] We tested this strategy during an autonomous traverse to the top of a rock-strewn hill at cite C. The rover began a distance from the base and traveled forward in 50 m intervals, collecting panoramas and workspace imagery at each locale. The traverse visited five locales. The first three were situated on the approach to the base of the hill and the fourth was located half-way to the top.

[83] To an untrained human observer at the site the first three locales all appeared alike, with occasional patches of white sediment and few significant rocks. The hillside terrain at the fourth locale was different, however—here the sediment contained some large dark rocks along with many small light-toned rocks. The light-toned material was no longer present at the fifth locale, but the peak of the hill was covered with large grey rocks. Figure 12 shows some of the collected data. Row A is an illustration corresponding to the basic human interpretation with dots representing material of different albedos. Row B shows sections of the 95-image panoramas collected at each locale. Row C shows underbody “workspace” imagery. Underbody images were not analyzed but they do provide some intuition into terrain conditions at each locale.

[84] We detected rocks in the five panoramas and analyzed them using a variety of different classification schemes. The first classification method favored manually chosen interval classes: an albedo feature interval to account for the light-toned material, a second interval based on rock size, and an outlier cluster to account for everything else. These distribution histograms appear in row D. The clusters show a general trend toward an increase in the number of rocks as one approaches the peak together with a corresponding drop in the proportion of light-toned material. Note that in both unsupervised and supervised cases, a larger number of light-toned rocks appear in locales I-III than the ground-truth records suggest. This is due to patches of light-toned sediment (visible the first three panoramas of row B) that can be mistaken for rocks.

[85] The second classification used a completely automatic model. Five multivariate Gaussian clusters were initialized to random data points. Then EM clustering iteratively adjusted the clusters' parameters to converge on values that maximized their likelihood with respect to the entire data set. The resulting distribution histograms appear in Figure 12 row E. Three color-channels and size features were used in clustering, but clusters divided mainly along the intensity axis. Because of this the legend provides mean intensity information to distinguish the clusters. This classification succeeded in detecting the general trend in number and types of rocks—fewer rocks in locales I-III, a change in surface composition as the rover climbed the hill, and a profusion of dark rocks at locale V.

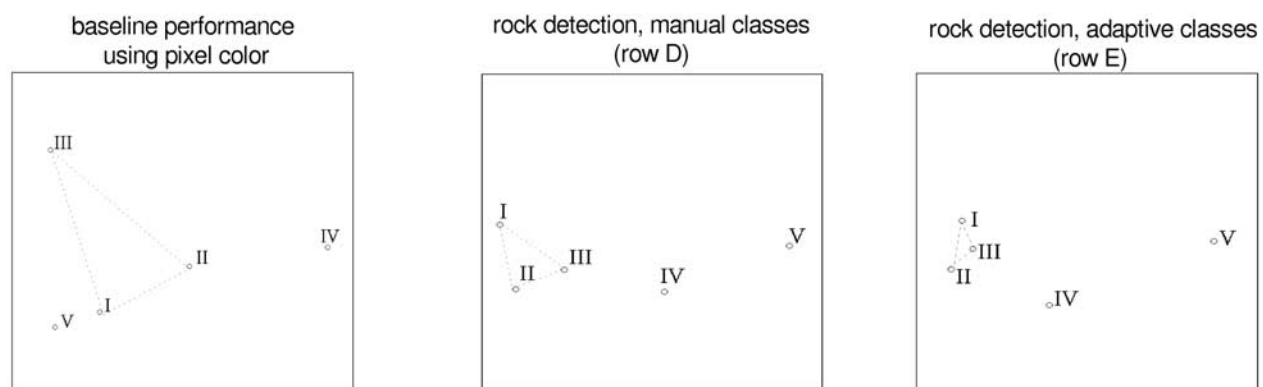


**Figure 12.** Five locales during a 200 m traverse. (a) Human interpretation. (b) Portion of the color panorama. (c) Underbody "workspace" imagery. (d) Manually fixed interval classes. (e) Classes from unsupervised clustering.

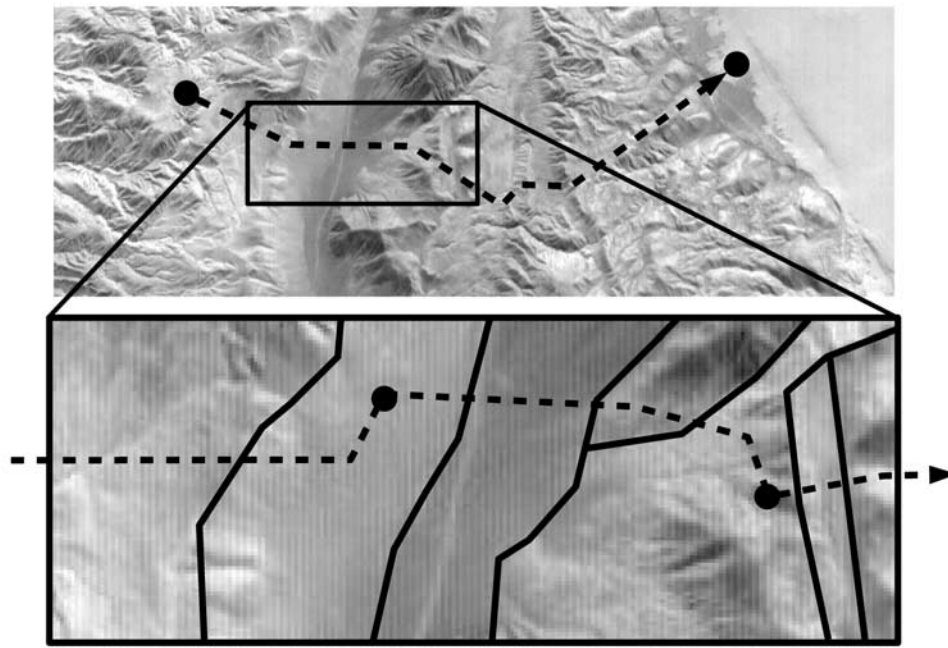
[86] The histograms in row E accurately reflect the similarity of locales I to III and the relative differences of locales IV and V. However, comparing these histograms visually is difficult. Figure 13 provides a simpler visual representation based on Principal Component Analysis (PCA). It shows the normalized histograms projected onto their first two principal components, yielding a 2D space that best preserves the variance of the histogram data. The figure compares three strategies for producing geologic signatures. First, a baseline control procedure used a simple histogram of pixel color values in the panorama as a proxy for geologic content. These color features fared worst; the

relationship between color pixels from different locales had little to do with the material present at the surface. This inaccuracy was highlighted by the first three locales, which were similar geologically but whose signatures were widely separated in the PCA projection. The distribution of colors was influenced by many factors—such as features on the horizon or lighting changes—that had little to do with the locale's geology.

[87] Detecting rocks in images provided better geologic signatures. The PCA plot of manual categories suggested a linear gradient of change between locales I and V. The ground truth geologic change is sudden, however, so this



**Figure 13.** PCA projection of class distribution histograms from Figure 12 helps to visualize differences between locales. Rock detection with automatic (adaptive) classification offers the best overall fidelity; it reflects the similarity of locales I-III and the unique surface compositions of locale IV and locale V.



**Figure 14.** Representative sampling strategy at multiple scales.

signature was still slightly misleading. The error could have occurred due to the bias of an overly rigid model. Much of the variation between locales occurred within the most general class and was thus invisible to the classification. Rock detection with unsupervised clustering gave the best result, matching the human interpretation that locales I–III were similar while locales IV and V were each different from all others.

## 5. Tactical Replanning to Efficiently Map the Distribution of Life

[88] One of the ultimate goals of rover exploration is to build maps that relate to models of the environment. For instance, in the Atacama, surface habitats for lichens and bacteria can be created by terrain features that locally modify air flow and insolation. The interaction of these variables with the presence of life is poorly understood. We propose a form of representative sampling called “sampling by regions” to efficiently answer this type of question.

[89] Under the sampling by regions strategy, regions with homogeneous properties are identified in orbital imagery. The rover is used to characterize aspects of each region that cannot be measured from orbit. In the Atacama, orbital data can be used to study frequency of cloud cover, landforms that control airflow, and average slope, which affects insolation and wind exposure. One can identify local regions that are homogeneous with respect to these properties, and then address the properties’ relationship to viable habitats by using the rover to characterize the presence or absence of life on a per-region basis.

### 5.1. Related Work in Planning and Execution

[90] There has been considerable research on planning systems for planetary surface rovers [Estlin *et al.*, 2002; Urmson *et al.*, 2003; Tompkins, 2005], achieving notable

success with the MAPGEN and GESTALT planners used by the Mars Exploration Rovers [Bresina *et al.*, 2005; Maimone *et al.*, 2006]. Our discussion will focus narrowly on controllers that either use probabilistic planning or have the ability to react to science observations collected on the fly.

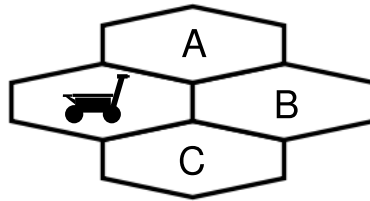
[91] The Antarctic meteorite search [Pedersen, 2001] (see also section 4.4) selected sensing actions greedily in order to reduce the uncertainty in rock/meteorite classification according to a Bayesian network model. Although its action selection was myopic, this system was notable for its principled use of a probabilistic model to handle uncertainty.

[92] Estlin *et al.* [2003] developed a planner for inserting trigger-response opportunistic science actions into a realistic rover command sequence. Their CASPER planner is sophisticated enough to reason about such issues as detailed path planning, instrument warm-up periods, and instantaneous power constraints that prevent rover subsystems from operating simultaneously. It successfully inserted rock measurement actions into a rover traverse as part of an integrated demonstration of the OASIS system in a controlled outdoor environment. The Autonomous Sciencecraft Experiment discussed in section 4.3 also used CASPER technology to insert opportunistic science actions [Chien *et al.*, 2003].

[93] Dearden *et al.* [2003] generated contingent plans that were robust with respect to uncertainty about the resources that actions would consume. One could use the same contingent plan representation to build in slack for opportunistic science.

[94] Smith [2004] developed a system for optimally dropping and reordering actions when rover resources are oversubscribed. Pedersen *et al.* [2006] described an integrated demonstration using Smith’s oversubscription planner to insert new science tasks on the fly as they were requested by the human members of a human-robot team.





**Figure 15.** (left) Actions available to the rover. (right) The Zoë rover at the test site in Pittsburgh.

[95] The LITA operational strategy of large-area surveys inspired a new onboard planning system. We aim to account for both scientists' plan constraints and enhance mapping efficiency by making use of rover observations on the fly.

## 5.2. Mapping Scenario

[96] To operationalize sampling by regions, one can describe the environment using the different size scales shown in Figure 14. At large scale (kilometers, upper map), scientists designate a path based on long-term mission goals. At small scale (hundreds of meters, lower map), the rover's path should be optimized to travel through regions relevant to the modeling objectives. In the figure, region boundaries are indicated with solid lines and the rover's path is marked with a dashed line.

[97] Within the rover's field of view (tens of meters) in the small scale map, the rover should scan the environment and react appropriately if it detects cues associated with life, such as visible plants or signs of gully erosion. This reaction might include taking a detour toward the feature and gathering additional data to confirm the presence of life. In the figure, detected features are marked with black circles; note that the rover has adjusted its path to visit them.

[98] We have developed an onboard system that intelligently controls rover motion and sensing at small scale, while respecting the constraints of the large scale plan specified by scientists. The first step is off-board probabilistic planning based on large-scale goals and a region map provided by scientists. This generates a robust policy that specifies how to react to different potential science instrument readings during execution. During traverse the rover refers to this policy to choose appropriate actions based on the actual instrument readings. In effect, it continuously replans the remainder of the traverse based on what it sees.

[99] The tactical replanning system attempts to follow the most efficient exploration strategy. In order to define optimality, where there are multiple interacting goals, each goal must be assigned a relative priority that allows the system to perform appropriate trade-offs. Our scenario encodes these priorities by specifying a per-region reward for finding evidence of life and penalties for taking time-consuming actions.

[100] In addition, one must specify how to deal with uncertainty. If a plan includes searching a region for life, and it is unknown whether the region contains life, one cannot evaluate in advance how successful the plan will be. We take a decision-theoretic perspective toward uncertainty. Uncertain aspects of the environment model and uncertain outcomes of rover actions are modeled probabilistically, and the optimal plan is the one that has the highest expected or "average" value.

## 5.3. Tactical Replanning Experimental Procedure

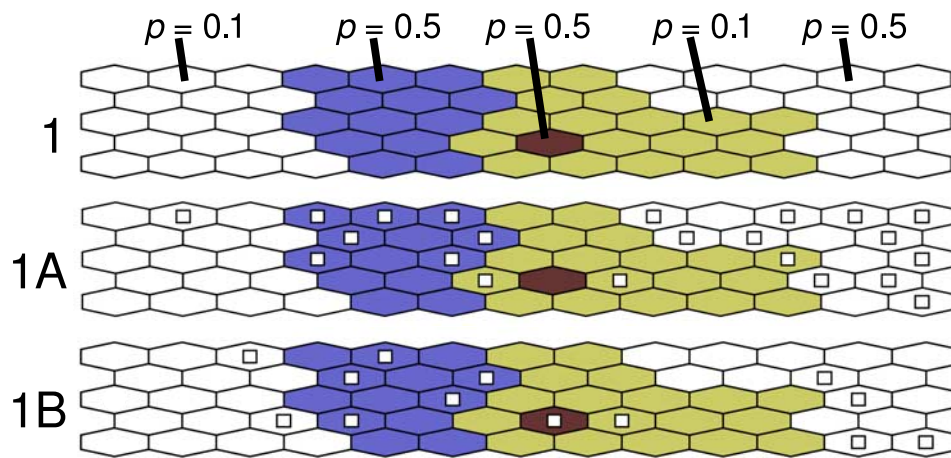
[101] We evaluated our tactical replanning system onboard the Zoë rover in a controlled outdoor test site in Pittsburgh (Figure 15, right). This allowed us to simplify several aspects of the problem. First, since there were no appropriately distributed natural features in our test area, we used simple artificial targets (10 cm squares of white posterboard) as a stand-in for evidence of life.

[102] Second, rather than generating a region map based on satellite images of the test area, we generated an arbitrary region map by hand and then modified the test area to match the map. To be more precise, since the region map only specifies the likelihood of life in each cell, we randomly generated multiple target layouts for each region map, using the specified likelihoods.

[103] Figure 16 shows one region map and two corresponding randomly drawn target layouts. Regions are indicated as groups of cells with the same shading. Each region is marked with the probability that cells in that region will contain evidence of life. White squares in the target layouts indicate the locations of artificial targets.

[104] In keeping with the overall scenario, the region map represents a segment of a long traverse. The rover starts at the left side of the map and must eventually exit from the right side in order to conform to the large scale plan.

[105] Figure 15 (left) helps to explain the actions available to the rover at each step of execution. In a single action, the rover could either (1) scan all three of the marked cells, returning a noisy signal as to whether they contain life, or (2) perform a simple move or sampling move to any one of the marked cells. Sampling moves differ from simple moves in that they cause the rover to take additional detailed



**Figure 16.** Prior map 1 and corresponding randomly drawn target layouts 1A and 1B.

measurements as it enters the new cell. They are intended to confirm the presence of life.

[106] The scan action returned one of three possible observations for the likelihood of life in each forward cell, interpreted roughly as “negative”, “maybe”, and “positive”. In onboard testing, these corresponded to different confidence levels from the detection routine searching for artificial markers, which due to the small size of the markers reported false positives about 20% of the time and false negatives about 10%. The sensor noise parameters used in the planning model were learned from a training set that included detection routine outputs and ground truth labels gathered over several runs in the testing environment.

[107] The planning objective was to maximize the “efficiency score” for the task, which tallied rewards and costs. The robot received a per-region reward: +5 points if the robot entered the region, +20 points if it passed through a life-containing cell in the region, or +50 points if it performed a sampling move into a life-containing cell in the region. Each action incurred a cost: −1 point for each move, and −5 points for each scan or sampling move. Thus the rover must find confirmed evidence of life in as many regions as possible, while minimizing the number of detours, scans, and sampling moves.

[108] We evaluated three planners on the life survey task. Under the blind planner, the rover simply moved to the right in a straight line, always using sampling moves. The blind planner would confirm the presence of life only if it was found on the straight-line path.

[109] Under the reactive planner, the rover followed a set of simple hand-generated rules designed to efficiently confirm the presence of life through combined use of scanning and sampling. It moved forward through the map, performing a scan action after each move, and detoured to take a sampling move if life was detected in a scan. When life was not detected, the reactive planner tried to stay on a preplanned path that was optimized to pass by areas likely to contain life.

[110] Under the probabilistic planner, actions were chosen in order to explicitly maximize the efficiency score defined earlier. The probabilistic planner was not limited to using a preconceived exploration strategy. Instead, it attempted to find the best strategy for any particular map.

[111] The probabilistic planner drew on a large body of research in the artificial intelligence community; we modeled the system as a partially observable Markov decision process, and we found a near-optimal plan using heuristic search and value iteration techniques [Smith and Simmons, 2005]. Given a prior map, the planner generated a complete plan with contingent branches for different possible observations during 20 minutes of off-board planning on a 3.2 GHz Pentium-4 processor, and this plan was transmitted to the rover for execution.

[112] The reactive and probabilistic planners were each evaluated on 20 runs through the test course; there were 2 prior maps, times 2 randomly drawn target layouts per map, times 5 runs per target layout. The blind planner could be evaluated on the same target layouts in simulation since its actions do not depend on uncertain sensor observations.

[113] Results are shown in Table 2. For each planner, we report average values over the 20 runs. “Search actions” gives the number of scan and sampling move actions used per run (smaller values are better). “Regions confirmed” gives the number of regions in which the presence of life was confirmed with a sampling move action (higher values are better). Finally, “efficiency score” is the combined efficiency metric that the planners were trying to optimize (higher values are better).

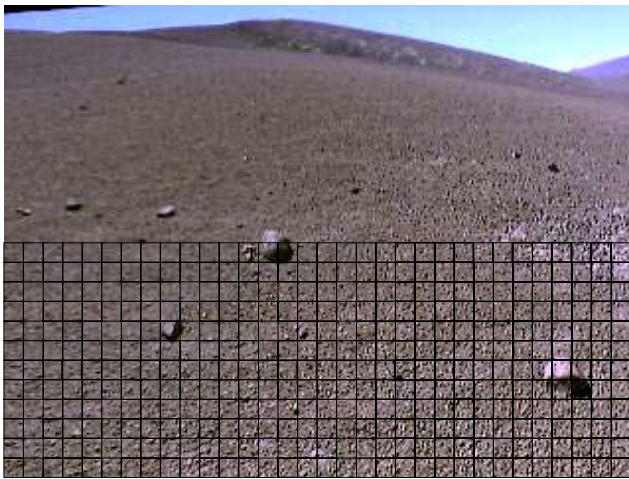
[114] The probabilistic planner performed best in terms of search actions and overall efficiency. The reactive planner confirmed the presence of life in more regions, but at the cost of far more search time than the other planners, which would leave less time for science on later legs of the traverse.

[115] These results demonstrate that rover exploration efficiency in a mapping task can be improved by reasoning about observations obtained en route. Moreover, general-purpose probabilistic planning techniques are computation-

**Table 2.** Tactical Replanning Experimental Performance

Planner	Search Actions	Regions Confirmed	Efficiency Score
Blind	12.0	2.5	68
Reactive	20.0	3.4	61
Probabilistic	7.5	3.0	113





**Figure 17.** A navigation image with foreground grid overlay. We extract numerical color and texture attributes from each grid square.

ally feasible in this domain, and were shown to outperform a reactive planner relying on hand-tuned heuristic rules. In future work, we would like to remove some of the simplifying assumptions of our test scenario and study rover mapping in a more realistic setting.

## 6. Detecting Novel Images and Geologic Unit Boundaries in Image Sequences

[116] During the third year of the LITA campaign scientists often commanded long sequences of navigation images. A navigation sequence consisted of  $320 \times 240$  pixel color images captured at two meter intervals from cameras mounted on the rover mast. They provided a  $60^\circ$  field of view of the terrain in front of the rover. Scientists exploited these sequences to gathering dense survey imagery and for gaining situational awareness of the rover traverse area. We hypothesized that navigation images were also amenable for autonomous analysis.

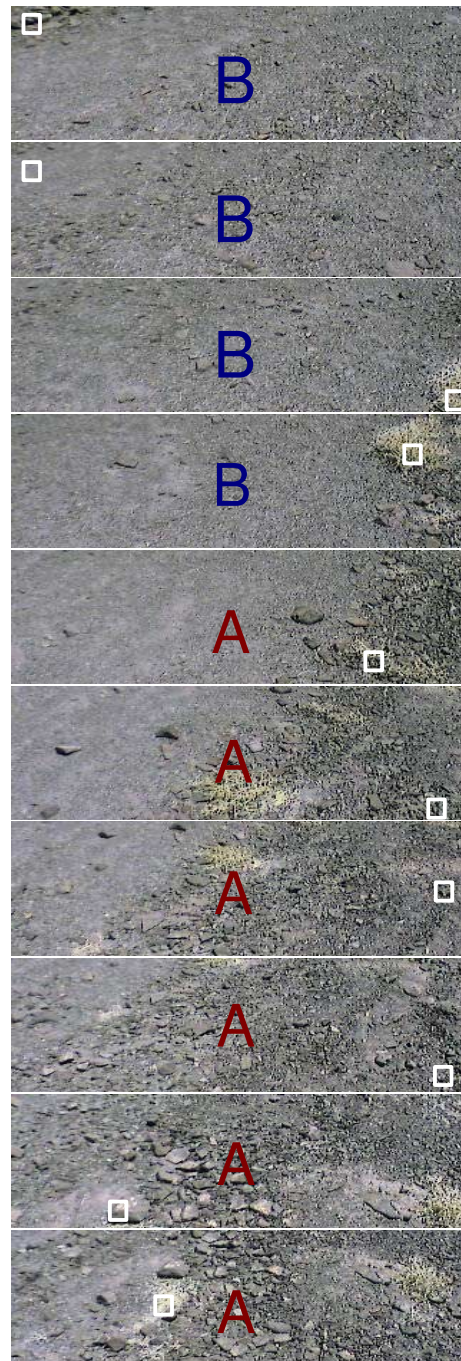
[117] The navigation sequences themselves often contained long stretches of homogeneous terrain with just a few outstanding or interesting transition frames. Onboard analysis was a prime candidate for reducing these data sets by identifying novel or interesting frames. Another unique aspect of navigation sequences was their scale—they could be as long as a multiple-kilometer traverse itself (Figure 4). Therefore trends in navigation sequences could theoretically assist with the identification of geologic boundaries. Their large scale permitted the integration of orbital data with surface information to further improve autonomous mapping.

[118] Context-sensitive models were used to analyze navigation sequences. As in the life survey experiment above, these models considered not only the feature itself but also its location. They identified novel features by highlighting anomalies that were unexpected in the context of the local environment. These local dependencies between neighboring sample points were used to localize geologic boundaries. Tests on navigation imagery from the LITA field season suggested that context-sensitive data analysis using a hidden Markov model [Rabiner, 1989] provided

novelty detection performance superior to periodic selection strategies.

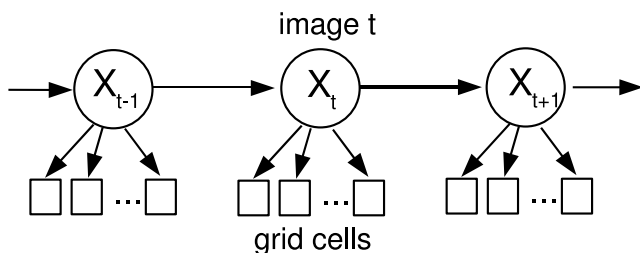
### 6.1. Novelty Detection in Navigation Sequences

[119] We evaluated novelty detection with a sequence of 1888 navigation images from a single site F traverse (LITA sol F20). We chose this sequence both for the long length of the traverse and the diversity of its terrain. This drive action began on a rocky hillside and transitioned to a basin covered in finer material.



**Figure 18.** Automatic labeling of a boundary in the environment. Letters correspond to the image class, while squares suggest novel locations for measurement.





**Figure 19.** The hidden Markov model used for novelty detection.

[120] The image frames in the navigation sequence had too poor a resolution for our rock detection algorithm. Instead we relied on simpler features as a proxy for geologic content. We split the foreground half of each image into a grid of 10-pixel cells; each cell constituted an independent sample of the imaging site (Figure 17). The system computed numerical attributes for the color and texture of each cell; a single image provided several hundred distinct samples in color-texture space. We color-normalized images using the “greyworld” strategy [Buchsbau, 1980] and used cells’ fractal dimension [Chaudhuri and Sarkar, 1995] as a simple texture measure.

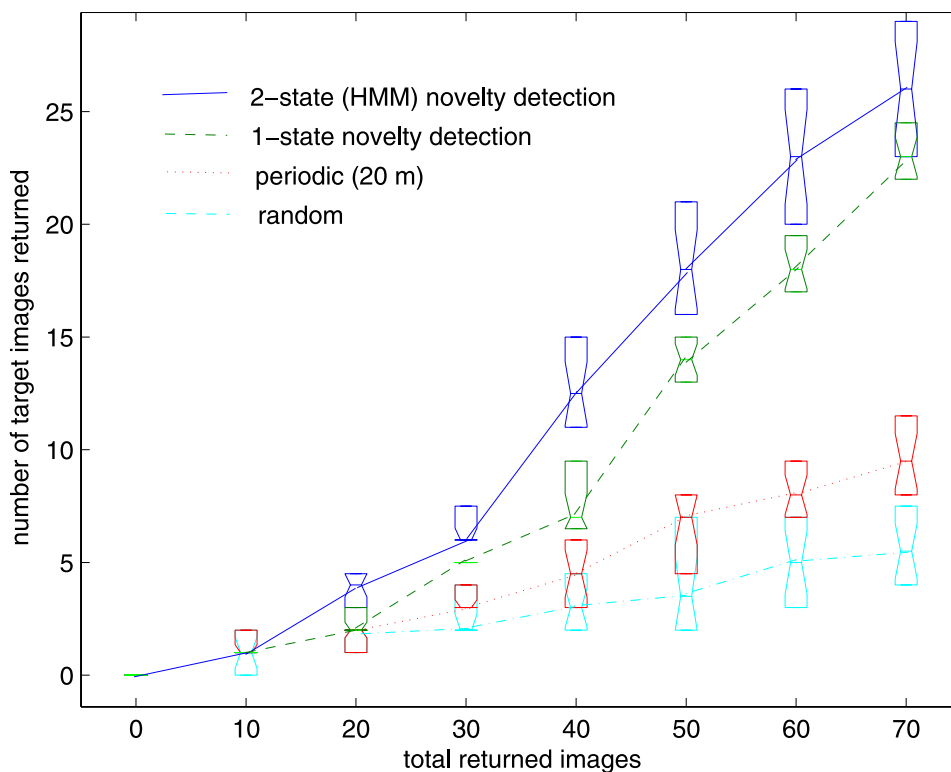
[121] The sequential structure of rover traverse imagery made it amenable to description by a hidden Markov model (HMM). The HMM used persistent “hidden states” to estimate the unobserved biologic or geologic characteristics of the rover’s environment. This permitted context-sensitive novelty detection; the rover could account for the influence

of neighboring observations when evaluating an image’s likelihood (Figure 18).

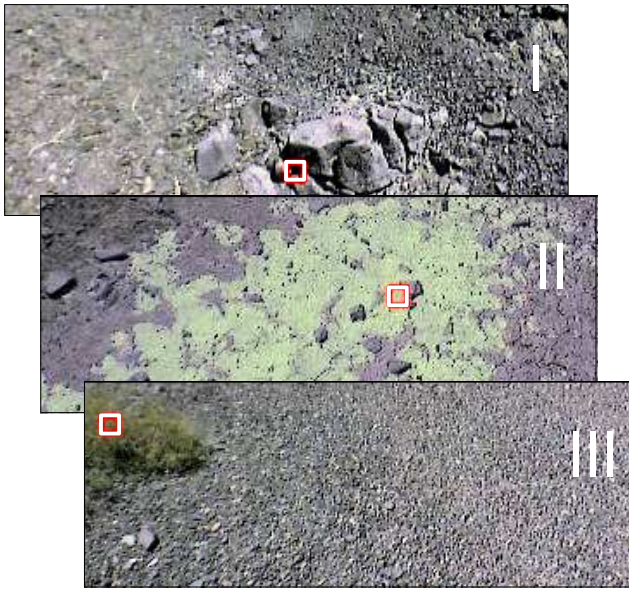
[122] A graphical representation of the HMM appears in Figure 19. The model quantified relationships between neighboring states (transition probabilities) and the likelihood of each state generating a particular observed grid cell (emission probabilities). Its chain structure reflected the persistence of environmental conditions; the current state influenced expectations for the next image’s content. Our implementation described observation probabilities with a tied continuous-density Gaussian mixture model [Belle-garda and Nahamoo, 1990]. The Baum-Welsh algorithm [Baum, 1972] assigned model parameters which could then be used to estimate the hidden state of any image. We estimated these parameters using the entire image sequence but one could also build models from partial data for adaptive sampling decisions during traverse.

[123] We evaluated the HMM system’s novelty detection performance for prioritizing images for transmission to Earth. Standard novelty detection strategies fit a statistical model to the entire data set and select low-likelihood outliers for downlink. We hypothesized that by representing the spatial structure of the environment, the HMM could improve performance by recognizing images that were unlikely in the context of local observations. We performed a relative performance evaluation with a set of rare image features that were unambiguous to identify. We used plants (appearing in 132 images) rocks with a major axis longer than 30 pixels (31 images), and the rover shadow (28 images).

[124] The novelty detection methods prioritized all sequence frames in order of decreasing novelty (increasing



**Figure 20.** Selective return using image analysis outperforms periodic and random return strategies for novelty detection. The context-sensitive HMM performs best of all.



**Figure 21.** Novel frames identified from the traverse. Squares indicate the single most novel grid cell.

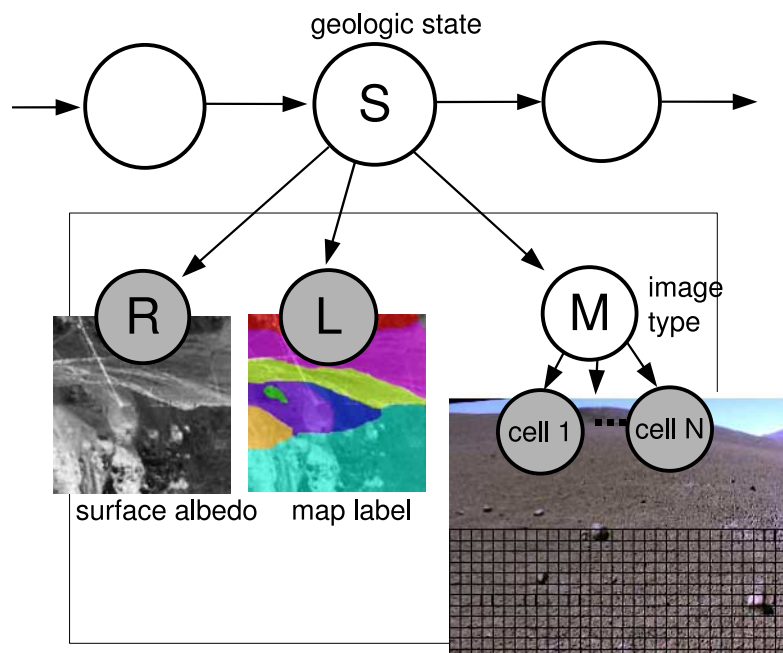
likelihood). The number of novel targets chosen for various data allowances indicated how closely different models' likelihood estimates aligned with the rare image features. This number produced a relative performance score. We used random model initializations to generate 20 trials simulating two different novelty detection strategies: a two-state HMM with three mixture components, and a single-state three-component mixture model. The one-state model was not context-sensitive; it was equivalent to a stationary Gaussian mixture model that ignored each image's sequence position and treated all data products as independent.

[125] Figure 20 shows the results of the experiment with boxes illustrating the middle data quartile and notches the 95% confidence interval. A random data return policy provided a performance baseline. Both novelty detection methods exhibited good selectivity for the chosen features. The context-sensitive HMMs provided the best overall performance for this data set. Figure 21 showed examples of novel image frames favored by the HMM, with an annotated square to indicate the most novel cell in each frame. These novel cells are possible targets for selective instrument placement.

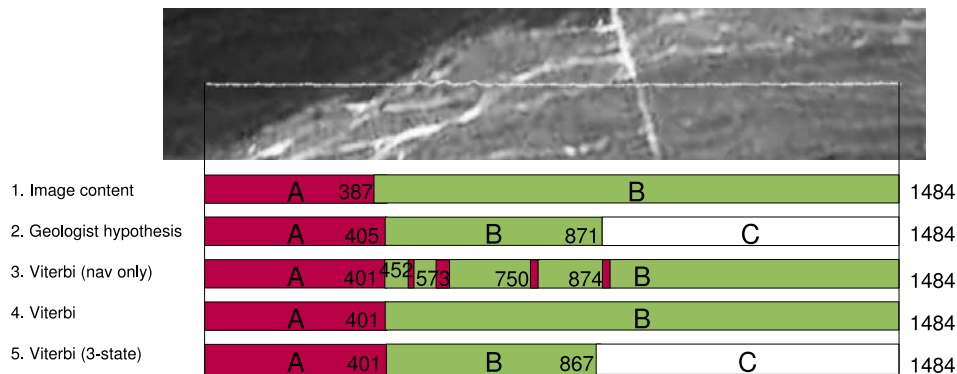
## 6.2. Identification of Geologic Units

[126] Long image sequences gathered during the LITA field campaign provided an opportunity to study automatic detection of geologic unit borders. For our purposes “geologic unit” refers simply to an area with a particular composition of visible surface material. Autonomous detection of geologic units could be used to trigger additional measurements at contact points or identify unexpected geology to prioritize for downlink. The experiment here used data collected during science operations at site F. Four traverse sequences from this site had crossed boundaries which project scientists had previously identified as probable contact points between units.

[127] The hidden state of the HMM provided a natural framework for mapping geologic state transitions during long-distance autonomous drives. Given a sequence of measurement sites, the Viterbi algorithm [Forney, 1973] identified the most likely sequence of hidden states. Figure 18 shows an example segment of the day 20 traverse labeled with two geologic states. On this particular boundary the rover crossed from a rocky berm onto a flat road. The Viterbi algorithm produces a state labeling most likely to produce the observed sequence of color and texture features in the images.



**Figure 22.** The hidden Markov model augmented with orbital data.



**Figure 23.** Traverse A.

[128] These Viterbi labelings rely exclusively on information collected by navigation cameras. To improve performance further we augmented the HMM model so that it tied together this surface data with orbital imagery and the predictions of a project scientist. The result was the multi-scale structure shown in Figure 22. We hypothesize a discrete hidden state variable reflecting the geologic state of the environment at each imaging location along the traverse. This geologic state probabilistically generated observed orbital data (in our case, the intensity value of the associated pixel from an ASTER near-IR data product). It also generated a geologist's map labeling, and an image type—a hidden variable describing the discrete class of image observed locally by the rover.

[129] We grouped the entire set of observed image cells into a lexicon of 20 basic types using K-means clustering. Different clusters in the lexicon approximately represent the different types of surface materials such as rocks, shadows, sediments of varying texture, and salt deposits. In addition to cell types, we included discrete “image types” that generated cells according to a multinomial distribution. Thus each image type referred to a specific proportion of cell types. One image type might generate cells that are half sandy and half rocky, while another might represent a surface covered entirely with big smooth rocks and their shadows. A single surface unit could in fact produce several image types. For example, the rover would observe different albedo/texture profiles from the sunlit and shaded sides of the same scene.

[130] In addition to navigation imagery the algorithm utilized ASTER orbital data. The ASTER imagery has a resolution of 30 m/pixel, and provided a Near-IR albedo value for each sampling location. We georectified ASTER images of the traverse region, blurred them to compensate for an average misregistration of about 50 m, and discretized the resulting albedo measurements to yield a second observed feature for each sample location. Finally, maps drawn by remote geologists provided an additional feature for each sampling site. These maps indicated hypothetical unit regions that could result in a different distribution of surface material. Adding them to the feature set suggested appropriate clusterings to the HMM based on the scientists' hypotheses. These labelings were merely hypothetical so we represented them numerically to the HMM as a smoothed distribution over possible labelings at each pixel.

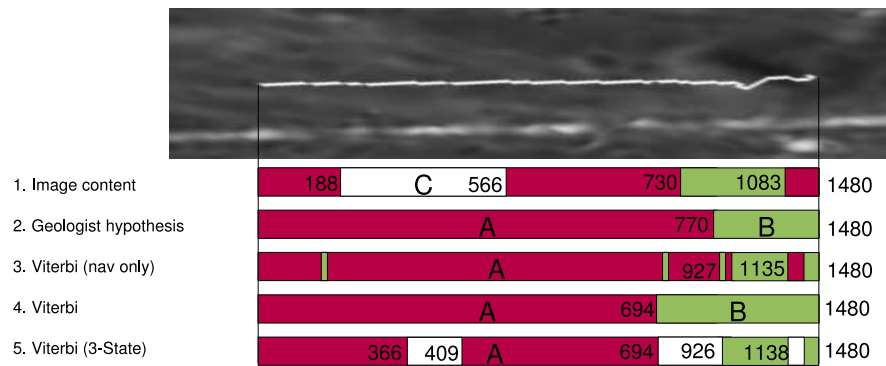
[131] Four traverses from the test data set crossed locations that had been hypothesized to be geologic boundaries. We used one of these traverses as a test set for tuning sensor processing and model parameters, and applied the HMM model to the remaining three traverses. We fit both cell and image type models on-line using the EM algorithm, fixing the number of image and cell types beforehand. We found that performance was robust to changes in these parameters so we left them at 5 and 10 respectively for the experiments below. After fitting the complete HMM to observed data we queried the model with the Viterbi algorithm [Forney, 1973]. This output the most probable hidden state parse of the traverse. We compared the resulting borders against ground-truth identification of the actual geologic border in imagery.

[132] Looking more closely at traverse A (Figure 23), the remote scientists hypothesized two geologic boundaries but the image set only provided obvious evidence for the first. In contrast, the 2-State parse unsurprisingly identifies this real border as the only boundary in the data set. Interestingly, the HMM also discovered a close approximation of this boundary autonomously using the sequence of navigation images alone and without any priming from orbital or human-labeled data. Forced to identify a third unit, it defaulted to the map label's suggestion.

[133] Traverse B (Figure 24) proved more difficult. While there was a perceptible change in material near image 414, all surface units in this traverse were highly heterogeneous. The second ground truth unit was in fact a mix of large rock-strewn and open areas. Moreover, the traverse was performed late in the day with rapidly changing light conditions. This further complicated the automated comparisons across images. The completely unsupervised model (using navigation imagery alone) failed to capture the intended structure. Including orbital data in the model improved the match to our ground truth parse, but the rover still falsely identified an unexpected new unit starting at image 1797. This was probably due to low light conditions at the end of the traverse.

[134] The terrain of traverse C (Figure 25) transitioned through more different surface compositions than the hypothesis labelings anticipated. The hypothesized geologic units implied a mostly homogeneous traverse with a single discrete transition at the base of a hill. In fact, the bare approach to the hill was punctuated by an unanticipated patch of cobbles. At image 683, the rover encountered the





**Figure 24.** Traverse B.

first evidence of the hill - a few rocks sourced from its peak. The rover started to climb the slope at image 724 through an increasing density of rocks. By image 790 it was deep in this hilltop field of rocks, which it navigated for the rest of the traverse. The only other interruption was a bare, sandy region beyond the peak beginning at image 1083.

[135] The HMM performed reasonably well on this traverse and captured most of the variation in surface material in each Viterbi parse. The differences between runs of the algorithm concerned the placement of transition points along the gradients and alternate interpretations of the material surrounding the peak. The fully autonomous “navigation image only” version treated the gradient of rocks as a series of discrete transitions, while the 3-state HMM used its extra descriptive power to model the slope as a separate geologic unit.

[136] The boundary-crossing traverses of LITA site F are a small but typical sample of boundaries scientists might encounter on a wide-area rover survey. Several aspects of the LITA data set are challenging. LITA scientists intentionally avoided rough terrain so most of these transitions concerned relatively minor variations in the composition of surface material. Moreover, the quality of image data available for processing was relatively poor—a  $320 \times 120$  pixel subimage from each location. Nevertheless the power of compounded evidence from multiple images permitted the HMM algorithm to identify meaningful boundaries in two out of three traverses. This was the case even the algorithm relied on surface imagery alone. This ability could assist planetary exploration robots in validating hypothesized boundaries, autonomously following geologic contacts for mapping purposes, or recognizing

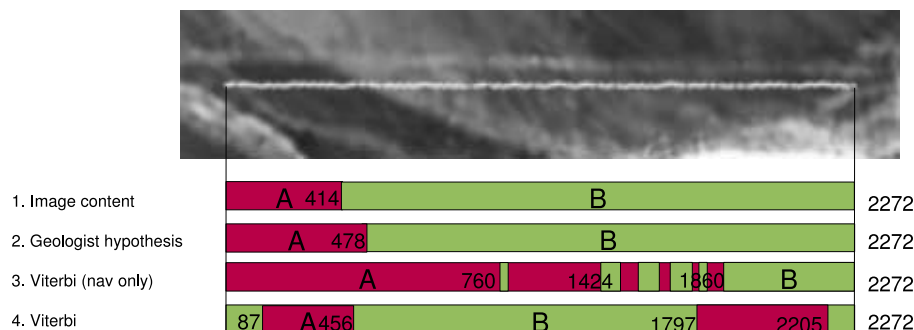
unexpected boundaries to trigger adaptive measurement actions.

## 7. Conclusion and Future Work

[137] The Life in the Atacama campaign provided an outstanding test bed for autonomous science by collecting massive data sets with a rover in Mars-analog terrain, employing a strategy of long traverses suited to responsive science on the fly, and driving innovations with its challenging science objectives. Science autonomy technology also returned benefits to the science investigation—by autonomously responding to evidence of life, the rover improved the efficiency of its life detection measurements and saved significant time and experimental materials (section 3).

[138] In the future, highly capable Mars rovers will drive longer distances and will face the communications bottlenecks inherent in over-the-horizon science. As we have shown, autonomous science can overcome these difficulties by adaptively planning data collection and return decisions that best match scientists’ exploration goals.

[139] The breadth of the LITA experiments underscores the fact that autonomous geology is not a single technology but rather a growing field of techniques and procedures. All four of the systems described in this paper were evaluated against simpler nonreactive exploration or data return strategies and showed significantly improved data quality. They used images collected during long rover traverses in a natural environment, with all the resulting challenges of varying lighting, partially buried rocks, and a broad spectrum of underlying geology. Table 3 summarizes the



**Figure 25.** Traverse C.

**Table 3.** Demonstrated Improvements in Data Quality

Autonomy	Result
Autonomously responding to evidence of life	90% increase in followup yield
Rock detection for selective data return	Scientists preferred 2 to 1 over periodic
Tactical replanning for mapping efficiency	Outperformed blind and reactive strategies
Novelty detection in image sequences	Outperformed stateless and periodic sampling

improvements in data return quality demonstrated by each of the experiments.

[140] In addition, there were several specific technical contributions:

[141] 1. The first demonstration of a science rover autonomously responding to the detection of life in the field (section 3).

[142] 2. Improved geologic classification using a machine learning framework that can cleanly integrate a variety of image processing techniques and data from multiple sensor types (section 4).

[143] 3. The first demonstration of autonomous mapping that takes both prior maps and rover science observations into account, and the first rover demonstration of a probabilistic planner that can reason about potential future science opportunities (section 5).

[144] 4. The first demonstration of automatic site characterization that fuses orbital and surface imagery (section 6).

[145] Our plans for future work extend in several directions. We plan to integrate aspects of the separate systems and remove simplifying assumptions, moving onboard science autonomy closer to flight readiness. We plan to apply our automated analysis techniques to existing Mars rover data sets. Finally, we are developing a theory of optimal control for robotic exploration tasks when the goal is to answer specific scientific questions.

[146] **Acknowledgments.** Scientific results and analysis were made possible by the efforts of the Life in the Atacama science team for which we are grateful. This research was supported by NASA under grants NNG0-4GB66G and NAG5-12890, Michael Meyer and Carl Pilcher, Program Scientists, and David Lavery, Program Executive.

## References

- Abrams, M. (2000), ASTER: Data products for the High Spatial Resolution Imager on NASA's EOS-AM1 platform, *Int. J. of Remote Sens.*, 21.
- Baum, L. E. (1972), An inequality and associated maximization technique in statistical estimation for probabilistic functions of Markov processes, *Inequalities*, 3, 1–8.
- Bellegarda, J. R., and D. Nahamoo (1990), Tied mixture continuous parameter modeling for speech recognition, *IEEE Trans. Acoust. Speech Signal Process.*, 38(12), 1033–1045.
- Bishop, C. M. (1995), *Neural Networks for Pattern Recognition*, Oxford Univ. Press, New York.
- Bornstein, B., and R. Castaño (2005), Creation and testing of an artificial neural network based carbonate detector for Mars rovers, paper presented at Aerospace Conference, Inst. of Electr. and Electron. Eng., Big Sky, Mont.
- Bresina, J. L., A. K. Jónsson, P. H. Morris, and K. Rajan (2005), Mixed-initiative planning in MAPGEN: Capabilities and shortcomings, paper presented at International Conference on Planning and Scheduling (ICAPS) Workshop on Mixed-Initiative Planning and Scheduling.
- Buchsbaum, G. (1980), A spatial processor model for object colour perception, *J. Franklin Inst.*
- Cabrol, N. A., et al. (2007), Life in the Atacama: Searching for life with rovers (science overview), *J. Geophys. Res.*, 112, G04S02, doi:10.1029/2006JG000298.
- Castaño, A., R. C. Anderson, R. Castaño, T. Estlin, and M. Judd (2004), Intensity-based rock detection for acquiring onboard rover science, *Proc. Lunar Planet. Sci. Conf.*
- Castaño, R., R. C. Anderson, T. Estlin, D. DeCoste, F. Fisher, D. Gaines, D. Mazzoni, and M. Judd (2003), Rover traverse science for increased mission science return, paper presented at Aerospace Conference, Inst. of Electr. and Electron. Eng., Big Sky, Mont.
- Chaudhuri, B. B., and N. Sarkar (1995), Texture segmentation using fractal dimension, *IEEE Trans. Pattern Anal. Mach. Intel.*, 2(1), 72–77.
- Chen, L., and P. Pu (2004), Survey of preference elicitation methods, *Tech. Rep. IC/2000467*, Swiss Fed. Inst. of Technol., Lausanne.
- Chien, S., et al. (2003), Autonomous science on the EO-1 mission, paper presented at International Symposium on Artificial Intelligence, Robotics, and Automation in Space (ISAIRAS), Nara, Japan.
- Chien, S., et al. (2005), Using autonomy flight software to improve science return on Earth Observing One, *J. Aerosp. Comput. Commun. Inf.*, 2, 196.
- Dearden, R., N. Meuleau, S. Ramakrishnan, D. Smith, and R. Washington (2003), Incremental contingency planning, paper presented at International Conference on Applied Planning and Scheduling (ICAPS) Workshop on Planning Under Uncertainty.
- Dempster, A. P., N. M. Laird, and D. B. Rubin (1977), Maximum likelihood from incomplete data via the EM algorithm, *J. R. Stat. Soc., Ser. B*, 39(1).
- Dial, G., H. Bowen, F. Gerlach, J. Grodecki, and R. Oleszczuk (2003), IKONOS satellite, imagery, and products, *Remote Sens. Environ.*, 88(1), 23–36.
- Dunlop, H. (2006), Automatic rock detection and classification in natural scenes, *Tech. Rep. CMU-RI-TR-06-40*, Robotics Inst., Carnegie Mellon Univ., Pittsburgh, Pa.
- Estlin, T., F. Fisher, D. Gaines, C. Chouinard, S. Schaffer, and I. Nesnas (2002), Continuous planning and execution for an autonomous rover, paper presented at International Workshop on Planning and Scheduling for Space, NASA, Houston, Tex.
- Estlin, T., R. Castaño, B. Anderson, D. Gaines, F. Fisher, and M. Judd (2003), Learning and planning for Mars rover science, paper presented at International Joint Conference on Artificial Intelligence (IJCAI) Workshop on Issues in Designing Physical Agents for Dynamic Real-Time Environments: World Modeling, Planning, Learning, and Communicating, Acapulco, Mex.
- Forney, G. D. (1973), The Viterbi algorithm, *Proc. IEEE*, 61(3), 268–278.
- Fox, J., R. C. No, and R. C. Anderson (2002), Onboard autonomous rock shape analysis for Mars rovers, paper presented at Aerospace Conference, Inst. of Electr. and Electron. Eng., Big Sky, Mont.
- Friedmann, E. I., and M. Galun (1974), Desert algae, lichens and fungi, in *Desert Biology*, edited by G. Brown, pp. 165–212, Academic, New York.
- Gilmore, M. S., R. Castaño, T. Mann, R. C. Anderson, E. D. Mjolsness, R. Manduchi, and R. S. Saunders (2000), Strategies for autonomous rovers at Mars, *J. Geophys. Res.*, 105(E12), 29,223–29,238.
- Gor, V., R. Castaño, R. Manduchi, R. C. Anderson, and E. Mjolsness (2000), Autonomous rock detection for Mars terrain, paper presented at Space Conference, Am. Inst. of Aeronaut. and Astronaut., Long Beach, Calif.
- Gulick, V. C., R. L. Morris, M. A. Ruzon, and T. L. Roush (2001), Autonomous image analysis during the 1999 Marsokhod rover field test, *J. Geophys. Res.*, 106(E4), 7745–7764.
- Maimone, M. W., J. Biesadecki, E. Tunstel, Y. Cheng, and P. C. Leger (2006), Surface navigation and mobility intelligence on the Mars Exploration Rovers, in *Intelligence for Space Robotics*, pp. 45–69, TSI Press, San Antonio, Tex.
- Pearl, J. (1988), *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann, New York.
- Pearlman, J., S. Carman, C. Segal, P. Jarecke, P. Clancy, and W. Browne (2001), Overview of the Hyperion Imaging Spectrometer for the NASA EO-1 mission, paper presented at International Geoscience and Remote Sensing Symposium (IGARSS).
- Pedersen, L. (2001), Autonomous characterization of unknown environments, paper presented at International Conference on Robotics and Automation (ICRA), Inst. of Electr. and Electron. Eng., Seoul, Korea.
- Pedersen, L. (2002), Science target assessment for Mars rover instrument deployment, paper presented at International Conference on Intelligent Robotics and Systems (IROS), Inst. of Electr. and Electron. Eng., Lausanne, Switzerland.

- Pedersen, L., M. D. Wagner, D. Apostolopoulos, and W. R. L. Whittaker (2001), Autonomous robotic meteorite identification in Antarctica, paper presented at International Conference on Robotics and Automation (ICRA), Inst. of Electr. and Electron. Eng., Seoul, Korea.
- Pedersen, L., et al. (2006), Field demonstration of surface human-robotic exploration activity, paper presented at Spring Symposium: To Boldly Go Where No Human-Robot Team Has Gone Before, Am. Assoc. for Artif. Intel., Stanford, Calif.
- Rabiner, R. L. (1989), A tutorial on hidden Markov models and selected applications in speech recognition, *Proc. IEEE*, 77(2), 257–286.
- Smith, D. E. (2004), Choosing objectives in oversubscription planning, paper presented at International Conference on Automated Planning and Scheduling (ICAPS).
- Smith, T., and R. Simmons (2005), Point-based POMDP algorithms: Improved analysis and implementation, paper presented at Conference on Uncertainty in Artificial Intelligence (UAI), Assoc. for Uncertainty in Artif. Intel., Edinburgh, Scotland.
- Smith, T., S. Niekum, D. Thompson, and D. Wettergreen (2005), Concepts for science autonomy during robotic traverse and survey, paper presented at Aerospace Conference, Inst. of Electr. and Electron. Eng., Big Sky, Mont.
- Thompson, D. R., S. Niekum, T. Smith, and D. Wettergreen (2005a), Automatic detection and classification of geological features of interest, paper presented at Aerospace Conference, Inst. of Electr. and Electron. Eng., Big Sky, Mont.
- Thompson, D. R., T. Smith, and D. Wettergreen (2005b), Data mining during rover traverse: From images to geologic signatures, paper presented at International Symposium on Artificial Intelligence, Robotics and Automation in Space (iSAIRAS), Munich, Germany.
- Tompkins, P. D. (2005), Mission-directed path planning for planetary rover exploration, Ph.D. thesis, Robotics Inst., Carnegie Mellon Univ., Pittsburgh, Pa.
- Urmson, C. P., R. G. Simmons, and I. Nesnas (2003), A generic framework for robotic navigation, paper presented at Aerospace Conference, Inst. of Electr. and Electron. Eng., Big Sky, Mont.
- Warren-Rhodes, K. A., K. L. Rhodes, S. B. Pointing, S. A. Ewing, D. C. Lacap, B. Gomez-Silva, R. Amundson, E. I. Friedmann, and C. P. McKay (2006), Hypolithic cyanobacteria, dry limit of photosynthesis, and microbial ecology in the hyperarid Atacama Desert, *Microb. Ecol.*
- Warren-Rhodes, K., et al. (2007), Searching for microbial life remotely: Satellite-to-rover habitat mapping in the Atacama Desert, Chile, *J. Geophys. Res.*, 112, G04S05, doi:10.1029/2006JG000283.
- Weinstein, S., et al. (2007), Application of pulsed-excitation fluorescence imager for daylight detection of sparse life in tests in the Atacama Desert, *J. Geophys. Res.*, doi:10.1029/2006JG000319, in press.

---

N. A. Cabrol and K. A. Warren-Rhodes, NASA Ames Research Center, MS 245-3, Moffett Field, CA 94035, USA.

T. Smith, D. R. Thompson, and D. S. Wettergreen, Robotics Institute, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15213, USA.

S. J. Weinstein, Molecular Biosensor and Imaging Center, Carnegie Mellon University, 5000 Forbes Ave., Pittsburgh, PA 15213, USA.